

THESES

présentées

A LA FACULTE DES SCIENCES DE L'UNIVERSITE DE PARIS

pour obtenir

LE GRADE DE DOCTEUR ES SCIENCES MATHÉMATIQUES

par

MOÏSE SIBONY

1ère thèse : *Sur l'approximation d'équations et inéquations aux dérivées partielles non linéaires de type monotone.*

2ème thèse : *Propositions données par la Faculté .*

Soutenues le 13 juin 1969 devant la Commission d'Examen

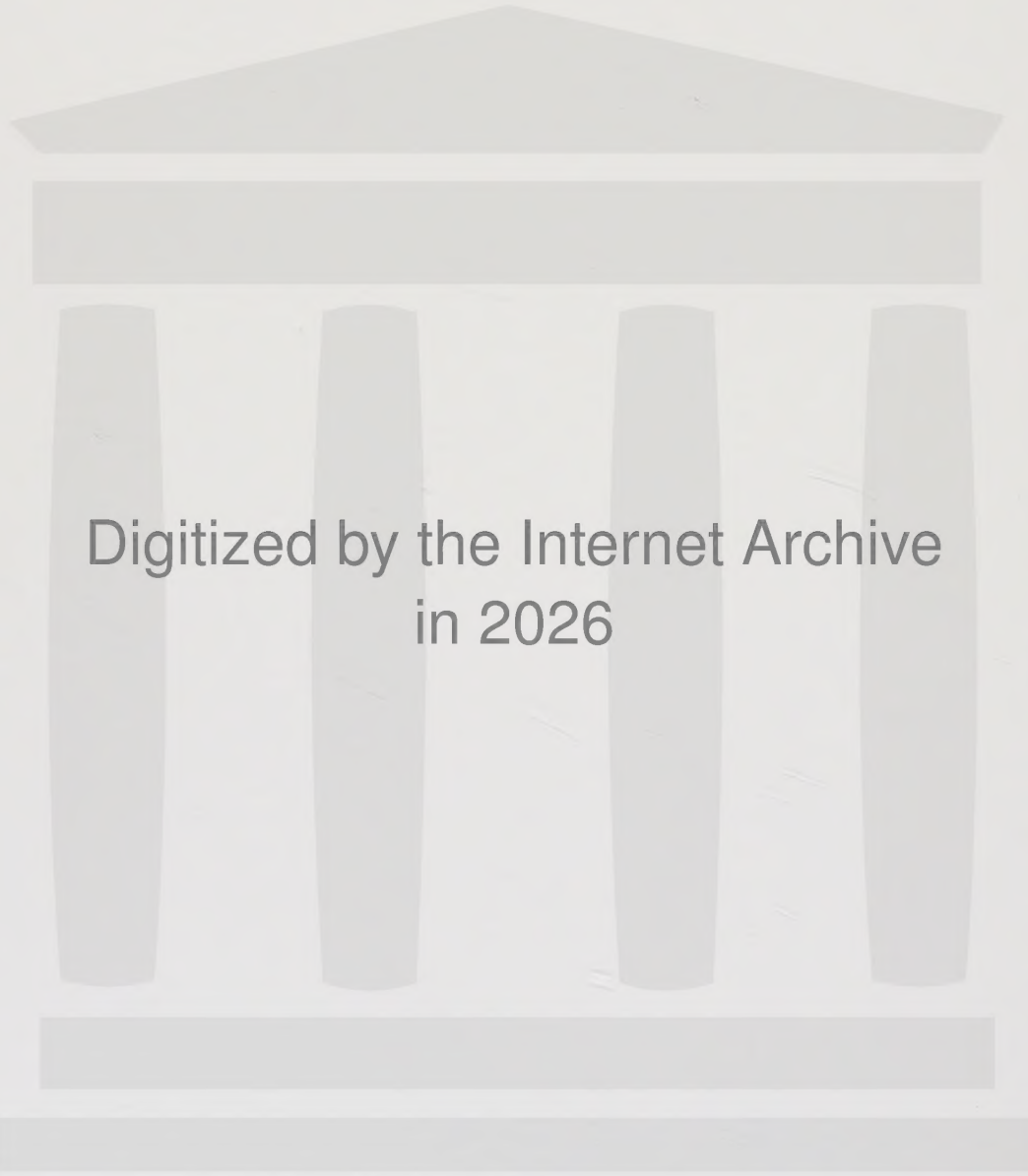
MM. L. SCHWARTZ

Président

J.L. LIONS

Examineurs

R. FORTET



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41551621_0029

THESES

présentées

A LA FACULTE DES SCIENCES DE L'UNIVERSITE DE PARIS

pour obtenir

LE GRADE DE DOCTEUR ES SCIENCES MATHÉMATIQUES

par

MOÏSE SIBONY

1ère thèse : *Sur l'approximation d'équations et inéquations aux dérivées partielles non linéaires de type monotone.*

2ème thèse : *Propositions données par la Faculté .*

Soutenues le 13 juin 1969 devant la Commission d'Examen

MM. L. SCHWARTZ

Président

J.L. LIONS

Examineurs

R. FORTET

232347

UNIVERSITY OF LEIDEN

LIBRARY OF THE UNIVERSITY OF LEIDEN

UNIVERSITY OF LEIDEN

LIBRARY OF THE UNIVERSITY OF LEIDEN

UNIVERSITY OF LEIDEN

LIBRARY OF THE UNIVERSITY OF LEIDEN



ב"ה

ויהי נעם אדני אלהינו עלינו ומעשה ידינו כוננה
עלינו ומעשה ידינו כוננהו. (Psaumes : 90, 17)

PREMIERE PARTIE

SUR L'APPROXIMATION D'EQUATIONS ET INEQUATIONS

AUX DERIVEES PARTIELLES NON LINEAIRES DE

TYPE MONOTONE

INTRODUCTION

Soit X un convexe fermé d'un Banach réel réflexif V , A un opérateur non nécessairement linéaire de V dans V' (dual de V) et f un élément de V' .

On cherche $u \in X$ tel que

$$(1) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X$$

où $(,)$ désigne la dualité entre V et V' .

Si $X = V$ le problème (1) se met sous la forme

$$(2) \quad Au = f$$

Par ailleurs, minimiser une fonctionnelle F sur le convexe X , est un problème qui peut se ramener à (1).

Des problèmes du genre (1) interviennent en mécanique, en théorie du contrôle optimal, et contiennent une classe assez large d'équations et inéquations aux dérivées partielles linéaires et non linéaires.

Notre travail consiste alors :

- 1) - A assurer l'existence et l'unicité de $u \in X$ solution de (1) et à élargir le cadre de (1) aux problèmes d'optimisation (chap I, I-1, I-2, I-3)
- 2) - Nous donnons ensuite des procédés systématiques pour l'approximation de u solution de (1) :

Le premier procédé est une régularisation elliptique (chap. I, I-4) de l'opérateur A (cf. [17]) élargissant en fait le cadre des hypothèses d'un procédé d'approximation de la solution de (1) à l'aide d'une méthode itérative, cf M. Sibony [22]. Ce même procédé est par ailleurs utilisé (chap. I, I-6) pour ramener le problème (1) à un problème analogue à (2).

Ensuite nous approchons à la manière d'Aubin (cf [1]) l'espace V par des espaces de dimension finie V_h ; ce qui revient à discrétiser le problème (chap I, I-5).

Il restera alors à donner une méthode constructive permettant de résoudre le nouveau problème en dimension finie.

La méthode de la mire (chap I, I-8) permet alors de résoudre directement le nouveau problème.

Dans [22] nous étudions des méthodes itératives, basées sur le calcul de la projection sur X , permettant de résoudre les problèmes (1) et (2).

Le chapitre II est consacré aux applications et aux résultats numériques. De nombreux exemples de ce chapitre sont traités numériquement dans [22] à l'aide de méthodes itératives.

Les deux parties de ce travail étant destinées à être publiées dans deux revues différentes nous avons préféré les laisser distinctes avec une bibliographie pour chacune.

Je suis profondément reconnaissant envers Monsieur le Professeur J.L. Lions qui a dirigé ce travail et dont les conseils et encouragements m'ont été très précieux.

Je remercie Monsieur le Professeur L. Schwartz d'avoir bien voulu accepter la présidence du Jury et Monsieur le Professeur R. Fortet de nous faire l'honneur d'y participer.

Enfin je remercie H. Brézis pour ses suggestions amicales.

CHAPITRE I

Théorèmes généraux

- I-1. Théorèmes d'existence et d'unicité
- I-2. Problème d'optimisation sur un convexe fermé
- I-3. Cas d'une fonctionnelle non différentiable
- I-4. Une méthode d'approximation par régularisation elliptique
- I-5. Une méthode d'approximation par discrétisation
- I-6. Méthode de pénalisation des contraintes
- I-7.. Méthode d'éclatement et pénalisation partielle des contraintes
- I-8. Optimisation de fonctionnelle sur un convexe fermé : méthodes directes de résolution

I - 1. Théorèmes d'existence et d'unicité

Soit V un espace de Banach sur $\mathbb{R}$ de dual V' . On désigne par $\|\cdot\|$ et $\|\cdot\|'$ les normes de V et V' respectivement. La forme bilinéaire $(\cdot, \cdot)$ désigne la dualité entre V et V' . Soit X un convexe fermé de V . On se donne un opérateur A non nécessairement linéaire de X dans V' ; et l'on cherche $u \in X$ vérifiant :

$$(1.1) \quad (Au, v-u) \geq (f, v-u) \quad \forall v \in X$$

pour f donné dans V' .

Le théorème 1.1 qui suit assure l'existence et l'unicité de la solution. Auparavant rappelons la terminologie :

Définition 1.1

On dit qu'un espace normé V est uniformément convexe si, $\forall \varepsilon$ tel que $0 < \varepsilon < 2$, $\exists \delta > 0$ tel que les relations : $\|u\| \leq 1$, $\|v\| \leq 1$ et $\|u-v\| \geq \varepsilon$, $u, v \in V$ impliquent $\|u+v\| \leq 2-\delta$.

Définition 1.2

1) On dit que $A : X \longrightarrow V'$ est monotone si

$$(Au - Av, u - v) \geq 0 \quad \forall u, v \in X$$

2) On dit que A est strictement monotone si

$$(Au - Av, u - v) > 0 \quad \text{pour tout } u, v \in X \text{ tels que } u \neq v.$$

Définition 1.3

On dit que $A : X \longrightarrow V'$ est hémicontinu si pour tout couple $u, v \in X$ l'application $t \in [0, 1] \longrightarrow (A((1-t)u + tv), u - v)$ est continue.

Théorème 1.1.

Soit V un Banach sur $\mathbb{R}$, réflexif ; X un convexe fermé de V contenant l'origine. On pose :

H1) A est un opérateur strictement monotone hémicontinu de X dans V' tel que

$$(1.2) \quad \lim_{\substack{\|u\| \rightarrow 0 \\ u \in X}} \frac{(Au, u)}{\|u\|} = +\infty \quad (\text{hypothèse sans objet si } X \text{ est borné})$$

H2) V est uniformément convexe et A est un opérateur monotone hémicontinu de

X dans V' vérifiant (1.2) ; de plus $(Au - Av, u - v) = 0$ entraîne $\|u\| = \|v\|$.

Alors, sous l'une des deux hypothèses H1) et H2) on a l'existence et l'unicité de la solution $u \in X$ de (1.1).

Démonstration.

1°) Supposons H1). Nous avons l'existence de la solution d'après un théorème de [9] cf aussi [4], [13], [14].

Montrons l'unicité de la solution. Soient u_1 et u_2 deux solutions de (1.1) distinctes. On a :

$$(1.3) \quad (Au_1, v-u_1) \geq (f, v-u_1) \quad \forall v \in X$$

$$(1.4) \quad (Au_2, v-u_2) \geq (f, v-u_2) \quad \forall v \in X$$

Posons $v = u_2$ dans (1.3) et $v = u_1$ dans (1.4). Par addition on a :

$$(Au_1 - Au_2, u_1 - u_2) \leq 0$$

et d'après la stricte monotonie de A on a $u_1 = u_2$.

2°) Montrons l'unicité sous l'hypothèse H2).

D'après la monotonie de A on a cette fois :

$$0 \leq (Au_1 - Au_2, u_1 - u_2) \leq 0. \text{ Donc } \|u_1\| = \|u_2\|$$

Soit S l'ensemble des solutions de (1.1). Comme dans [6] on montre que S est un ensemble convexe fermé. Comme V est uniformément convexe, donc strictement convexe ; alors S , convexe fermé situé sur une sphère est réduit à un point. D'où l'unicité.

Remarque 1.1

1°) L'hypothèse H3) suivante entraîne H1) et H2).

H3) V est uniformément convexe et A est un opérateur hémicontinu de X dans V' tel que

$$(1.5) \quad (Au - Av, u - v) \geq (\varphi(\|u\|) - \varphi(\|v\|)) (\|u\| - \|v\|)$$

où $\varphi = R_+ \longrightarrow R$ strictement croissante telle que

$$\lim_{r \rightarrow \infty} \varphi(r) = +\infty.$$

2°) Au lieu de (1.5) on fait l'hypothèse :

$$(1.6) \quad (Au - Av, u - v) \geq c \|u - v\|^2 \quad c > 0.$$

Sous cette hypothèse plus restrictive, d'autres démonstrations de l'existence et l'unicité de la solution u de (1.1) sont données dans [17] et [23] dans le cas où V est un Hilbert et A est un opérateur linéaire.

3°) On démontre facilement que toute solution $u \in X$ de (1.1) est solution de

$$(1.7) \quad (Av, v - u) \geq (f, v - u) \quad \forall v \in X$$

et réciproquement (cf [6]). Cette remarque est utile dans de nombreux passages à la limite.

Remarque 1.2

Si on fait $X = V$ le théorème 1.1 donne l'existence et l'unicité de $u \in V$ solution de l'équation

$$(1.8) \quad Au = f \text{ pour } f \text{ donné dans } V'.$$

Dans ce qui suit nous donnons un autre théorème d'existence et unicité de $u \in V$ solution de (1.8).

Posons

$$B_r = \{u \in V \mid \|u\| \leq r\}; S_r = \{u \in V \mid \|u\| = r\}, r \text{ réel } > 0$$

i) V est un Banach réflexif et A est un opérateur strictement monotone hémicontinu de B_r dans V' tel que $(Au, u) \geq 0 \quad \forall u \in S_r$

ii) V est un Banach uniformément convexe et A est un opérateur monotone hémicontinu de B_r dans V' vérifiant $(Au, u) \geq 0 \quad \forall u \in S_r$

De plus si $(Au - Av, u - v) = 0$ alors $\|u\| = \|v\|$. Nous avons alors le

Théorème 1.2

On suppose i) ou ii) alors, il existe $u \in B_r$ unique solution de

$$(1.9) \quad Au = 0$$

Démonstration

L'existence de u est montrée dans [3] sous l'hypothèse i)

Montrons l'unicité sous la même hypothèse : Soient u_1 et u_2 deux solutions de (1.9) avec $u_1 \neq u_2$. On a :

$$0 = (Au_1 - Au_2, u_1 - u_2); \text{ d'après la stricte monotonie de } A, \text{ on a } u_1 = u_2.$$

Sous l'hypothèse ii) on montre que l'ensemble S des solutions de (1.9) est un convexe fermé situé sur une sphère. Comme V est uniformément convexe ; alors S est réduit à un point.

Corollaire 1.1

Pour tout $f \in V'$, il existe $u \in V$ unique solution de

$$(1.10) \quad Au = f$$

sous l'une des deux hypothèses suivantes :

iii) V est un Banach réflexif et A est une application strictement monotone hémicontinue de V dans V' telle que

$$\lim_{\|u\| \rightarrow \infty} \frac{(Au, u)}{\|u\|} = +\infty \quad (u \in V).$$

iv) V est un Banach uniformément convexe et A est monotome hémicontinu tel que

$$\lim_{\|u\| \rightarrow \infty} \frac{(Au, u)}{\|u\|} = +\infty \quad (u \in V).$$

et si de plus $(Au - Av, u - v) = 0$ implique $\|u\| = \|v\|$.

Démonstration

Si $\lim_{\|u\| \rightarrow \infty} \frac{(Au, u)}{\|u\|} = +\infty$, alors $\exists r > 0$ tel que si $u \in V$ avec $\|u\| = r$ on a :

$$(Au, u) \geq r c(r) \text{ où } c \text{ est une application de } \mathbb{R}_+ \text{ dans } \mathbb{R}_+ \text{ telle que } \lim_{r \rightarrow \infty} c(r) = +\infty$$

En posant $\tilde{A}u = Au - f$ on a $(\tilde{A}u, u) \geq 0 \quad \forall u \in S_r$ pour r suffisamment grand. On applique alors le théorème 1.2.

I - 2. Problème de minimisation sur un convexe fermé.

Rappelons un résultat bien connu : Soit V un espace de Hilbert sur $\mathbb{R}$; $a(u, v)$ une forme bilinéaire continue sur V , V symétrique et vérifiant $a(u, u) \geq \alpha \|u\|^2$ avec $\alpha > 0$ pour tout $u \in V$. On se donne un ensemble X , convexe fermé de V . On sait que minimiser la forme :

(2.1) $J(v) = a(v, v) - 2(f, v)$, $v \in X$, pour f donné dans V' , équivaut à trouver la solution $u \in X$ de l'inéquation

(2.2) $a(u, v-u) \geq (f, v-u) \quad \forall v \in X$.

Le théorème 2.1 qui suit donne une généralisation non linéaire de ce résultat. Réciproquement nous allons montrer que tout problème d'optimisation (sur un convexe X) de la forme :

(2.3) $F(u) \leq F(v) \quad \forall v \in X$

où F est une fonctionnelle définie sur X , peut s'écrire sous forme d'inégalité variationnelle non linéaire.

Définition 2.1

Soit V un Banach sur $\mathbb{R}$; soit $u \in V$ et X un voisinage de u et F une fonctionnelle sur X ; on dit que F est dérivable au sens de Gateaux au point $u \in X$, si pour tout $v \in V$, l'application

$t \longrightarrow F(u + tv)$ est dérivable en $t = 0$. Soit $F'(u).v$ cette dérivée ; on suppose que l'application $v \longrightarrow F'(u).v$ est linéaire et continue de V dans $\mathbb{R}$.

On note :

$$(F'(u), v) = \lim_{t \rightarrow 0} \frac{F(u + tv) - F(u)}{t}.$$

Rappelons (cf [3]) pour la commodité du lecteur la

Proposition 2.1

Soit V un Banach réflexif sur $\mathbb{R}$; $F : V \longrightarrow \mathbb{R}$ dérivable au sens de Gateaux. Les conditions suivantes sont équivalentes :

i) F est convexe

ii) L'application $u \longrightarrow F'(u) \in V'$ est monotone sur V .

Démonstration

Posons $\psi(t) = F(tu + (1-t)v)$; $u, v \in V$.

i) $\implies$ ii). L'hypothèse entraîne que ψ est convexe et dérivable sur $[0, 1]$ et on a : $\psi'(t) = (F'(tu + (1-t)v), u-v)$.

Donc ψ' est croissante et par suite $\psi'(0) \leq \psi'(1)$.

D'où $(F'(u) - F'(v), u-v) \geq 0 \quad \forall u, v \in V$.

ii) $\Rightarrow$ i). L'hypothèse entraîne que ψ est dérivable avec ψ' croissante.
Donc ψ est convexe et donc F aussi.

Théorème 2.1

Soit V un Banach réflexif, F une fonctionnelle convexe sur X (convexe fermé de V), dérivable au sens de Gateaux.

Posons $A = F'$. Soit f donné dans V' . Alors toute solution de l'inégalité

$$(2.4) \quad (Au, v-u) \geq (f, v-u) \quad \forall v \in X$$

est solution de

$$(2.5) \quad J(u) \leq J(v) \quad \forall v \in X \quad \text{où } J(v) = F(v) - (f, v) \text{ et réciproquement.}$$

Démonstration.

1° Soit u solution de (2.5). Alors

$$F(u) - (f, u) \leq F(v) - (f, v) \quad \forall v \in X$$

Posons $v = u + t(w-u)$ $t > 0$, $w \in X$. On obtient :

$$\frac{F(u + t(w-u)) - F(u)}{t} \geq (f, w-u). \text{ En passant à la limite quand } t \rightarrow 0 \text{ on a :}$$

$$(F'(u), w-u) \geq (f, w-u). \text{ D'où}$$

$$(Au, w-u) \geq (f, w-u) \quad \forall w \in X.$$

2° Réciproquement, soit u solution de (2.4). D'après la monotonie de F' on a :

$$(F'(v), v-u) \geq (F'(u), v-u)$$

ce qui donne

$$(F'(v), v-u) \geq (f, v-u) \quad \forall v \in X.$$

Faisons $v = u + t(w-u)$ $t > 0$, $w \in X$. Il vient :

$$(2.6) \quad (F'(u + t(w-u)), w-u) \geq (f, w-u) \text{ pour tout } t \in [0, 1].$$

Posons :

$$h(t) = F(u + t(w-u)) - t(f, w-u) ; h \text{ est convexe car } F \text{ l'est.}$$

$$\text{Or } h'(t) = (F'(u + t(w-u)), w-u) - (f, w-u).$$

$$\text{Il résulte de (2.6) que } h'(t) > 0 \quad \forall t \in]0, 1].$$

$$\text{Donc } h(t) \text{ est croissante. En particulier } h(0) \leq h(1).$$

ce qui donne

$$F(u) \leq F(w) - (f, w-u). \text{ D'où } J(u) \leq J(w) \quad \forall w \in X.$$

Remarque 2.1

En particulier si V est un Hilbert et $a(u, v)$ est une forme bilinéaire symétrique et continue sur $V \times V$ telle que $a(v, v) \geq 0 \quad \forall v \in V$. Alors trouver $u \in X$ tel que $a(u, v-u) \geq (f, v-u) \quad \forall v \in X$

équivalait à minimiser dans X la fonctionnelle

$$J(v) = a(v, v) - 2(f, v).$$

Remarque 2.2

Comme le convexe X peut être non borné, si on fait $X = V$, le théorème 1.1 redonne l'existence et l'unicité de la solution $u \in V$ de l'équation

$$Au = f \text{ pour } f \text{ donné dans } V'.$$

Le théorème 2.1 dit alors que toute solution de l'équation $Au = f$ réalise le minimum dans V de la fonctionnelle $J(v) = F(v) - (f, v)$ et réciproquement.

Cette remarque est valable pour toute la suite.

Remarque 2.3

Soit $F : X \rightarrow \mathbb{R}$ convexe, dérivable au sens de Gateaux.

Si $A = F'$ vérifie les hypothèses du théorème 1.1 ; nous aurons l'existence et l'unicité de la solution $u \in X$ de l'inégalité $F(u) \leq F(v) \quad \forall v \in X$.

Nous allons montrer qu'on peut affaiblir ces hypothèses pour avoir un tel théorème d'existence et d'unicité de u solution de l'inégalité :

$$F(u) \leq F(v) \quad \forall v \in X.$$

Théorème 2.2

Soit V un Banach réflexif, X un convexe fermé de V . Soit F une application de X dans $]-\infty, +\infty]$ semi-continue inférieurement (s.c.i) pour la topologie induite par $\sigma(v, V')$. On suppose que F est non identique à $+\infty$ et :

$$(2.7) \lim_{\substack{\|u\| \rightarrow \infty \\ u \in X}} F(u) = +\infty \text{ (hypothèse sans objet si } X \text{ est borné)}$$

Alors

1° Il existe $u \in X$ tel que

$$(2.8) F(u) \leq F(v) \quad \forall v \in X$$

2° Si F est de plus strictement convexe ; on a l'unicité de u .

Démonstration.

1° Soit v_0 fixé dans X . Posons $K_{v_0} = \{u \mid u \in X ; F(u) \leq F(v_0)\}$

a) Montrons que (2.8) équivaut à

$$(2.9) F(u) \leq F(w) \quad \forall w \in K_{v_0}$$

Dans un sens l'assertion est évidente. Réciproquement supposons (2.9) et soit $v \in X$. Si $v \in K_{v_0}$ on a l'implication ; sinon on a : $F(v) > F(v_0)$. Donc

$$F(u) \leq F(v) \quad \forall v \in X.$$

b) Montrons que K_{V_0} est borné. Nous avons

(2.10) $F(u) \leq F(v_0)$. D'après (2.7), si K_{V_0} n'était pas borné alors $\exists u_n \in K_{V_0}$ avec $\|u_n\| \rightarrow \infty$ et le premier membre de (2.10) tendrait vers $+\infty$, ce qui est impossible.

c) K_{V_0} est donc borné dans V réflexif. Il est faiblement fermé car F est s.c.i pour la topologie faible donc faiblement compact. L'application $F : K_{V_0} \rightarrow \mathbb{R}$ est s.c.i pour la topologie faible. Donc $\exists u \in K_{V_0}$ qui réalise la borne inférieure de F .

2° Soient u_1 et u_2 deux solutions distinctes. Nous avons

$$F(tu_1 + (1-t)u_2) < t F(u_1) + (1-t) F(u_2) \quad \forall t \in]0, 1[$$

Ceci entraîne que $u_1 = u_2$.

Corollaire 2.1

Soit V un Banach réflexif ; X un convexe fermé de V .

Soit F une application de X dans $]-\infty ; +\infty]$ non identique à $+\infty$; s.c.i et strictement convexe. On suppose de plus que F est dérivable au sens de Gateaux et vérifie (2.7)' $\frac{F(u)}{\|u\|} \rightarrow \infty$ quand $\|u\| \rightarrow \infty$. Alors, il existe $u \in X$ unique tel que

$$(2.11) \quad (F'(u), v-u) \geq (f, v-u) \quad \forall v \in X \text{ pour } f \text{ donné dans } V'.$$

Démonstration

D'après le théorème 2.1 la relation (2.11) équivaut à chercher $u \in X$ tel que $J(u) \leq J(v) \quad \forall v \in X$ où $J(v) = F(v) - (f, v)$.

Le théorème 2.2 donne alors l'existence et l'unicité de la solution $u \in X$ de (2.11) ; car si F est convexe et s.c.i alors J est s.c.i pour la topologie faible induite par $\sigma(V, V')$.

Remarque 2.4

Dans le corollaire 2.1 on peut remplacer l'hypothèse : " F strictement convexe" par " F convexe et F' vérifiant :

$(F'(u) - F'(v), u-v) > 0$ pour $u \neq v$ ". Ces deux assertions sont d'ailleurs équivalentes.

I - 3. Cas où F est une fonctionnelle non différentiable.

Nous avons le

Théorème 3.1. Soit V un Banach réflexif sur R ; X un convexe fermé de V.

Soit $F : X \longrightarrow]-\infty, +\infty]$ non identique à $+\infty$. On suppose :

i) F s.c.i. et strictement convexe

ii) $\lim_{\substack{\|u\| \rightarrow \infty \\ u \in X}} F(u) = +\infty$. Alors

1°) Il existe $u \in X$ unique solution de

$$(3.1) \quad F(u) \leq F(v) \quad \forall v \in X$$

2°) Il existe un opérateur A de X dans V' monotone hémicontinu non nul et une fonctionnelle G sur X à valeur dans $]-\infty, +\infty]$ $G \neq F$ non identique à $+\infty$, s.c.i. et convexe tels que

toute solution de (3.1) est solution de

$$(3.2) \quad (-Au, v-u) \leq G(v) - G(u) \quad \forall v \in X$$

et réciproquement.

3°) L'inégalité (3.2) peut se mettre sous la forme : $u_1 \in X_1$

$$(3.3) \quad (A_1 u_1, v_1 - u_1) \geq (f_1, v_1 - u_1) \quad \forall v_1 \in X_1$$

avec $X_1 = \{v_1 \mid v_1 = (v, \lambda) \in X \times R ; G(v) \leq \lambda\}$

$$A_1 u_1 = (Au, 0)$$

$$f_1 = (0, -1)$$

Auparavant démontrons le

Lemme 3.1

Soient F_1 et F_2 deux fonctionnelles convexes sur X. On suppose F_1 dérivable au sens de Gateaux. Alors toute solution de

$$(3.4) \quad F(u) \leq F(v) \quad \forall v \in X \text{ avec } F = F_1 + F_2$$

est solution de

$$(3.5) \quad (-F'_1(u), v-u) \leq F_2(v) - F_2(u) \quad \forall v \in X \text{ et réciproquement.}$$

Démonstration

a) En posant $v = u + t(w-u)$ $t > 0, w \in X$, la relation (3.4) entraîne :

$$\frac{F_1(u) - F_1(u + t(w-u))}{t} \leq \frac{F_2(u + t(w-u)) - F_2(u)}{t}$$

comme F_2 est convexe on a :

$$\frac{F_1(u) - F_1(u + t(w-u))}{t} \leq F_2(w) - F_2(u)$$

En passant à la limite quant $t \rightarrow 0$ on a :

$$(-F'_1(u), w-u) \leq F_2(w) - F_2(u) \quad \forall w \in X. \text{ D'où (3.5).}$$

b) Réciproquement supposons (3.5) ; nous avons alors :

$$F(u) - F(v) = F_1(u) + F_2(u) - F_1(v) - F_2(v) + (-F'_1(u), v-u) + (F'_1(u), v-u)$$

D'après (3.5) on a :

$$F(u) - F(v) \leq F_1(u) - F_1(v) + (F'_1(u), v-u)$$

or si F_1 est dérivable on a :

$$(F'_1(u), v-u) \leq F_1(v) - F_1(u). \text{ ce qui donne :}$$

$$F(u) - F(v) \leq 0. \text{ D'où } F(u) \leq F(v) \quad \forall v \in X.$$

Démonstration du théorème 3.1

1°) Le théorème 2.2. assure l'existence et l'unicité de la solution $u \in X$ de (3.1).

2°) Posons $F = F_0 + F - F_0 = F_0 + G$ avec F_0 continue dérivable au sens de Gateaux et telle que $G = F - F_0$ soit convexe. (on pourra prendre F_0 linéaire). Posons $A = F'_0$. Nous avons alors, G s.c.i. et convexe. D'après le lemme 3.1, toute solution de (3.1) est solution de (3.2) et réciproquement.

3°) U. MOSCO a montré dans [19] qu'on pouvait mettre toute inégalité de la forme (3.6) $(f - Au, v-u) \leq G(v) - G(u) \quad \forall v \in X$ sous la forme : $u_1 \in X_1$.

$$(3.7) \quad (A_1 u_1, v_1 - u_1) \geq (f_1, v_1 - u_1) \quad \forall v_1 \in X_1$$

avec $v_1 = (v, \lambda) \in X \times \mathbb{R}$

$$A_1 u_1 = A_1(u, \bar{\lambda}) = (Au, 0)$$

$$f_1 = (f, -1)$$

$$\text{et } X_1 = \{v_1 \mid v_1 = (v, \lambda) \in X \times \mathbb{R} ; G(v) \leq \lambda\}$$

D'où le théorème.

Remarque 3.1

D'après la remarque 1.1, 3°), et ce qui précède on démontre facilement les assertions suivantes :

1°) Si F est une fonctionnelle sur X différentiable au sens de Gateaux alors toute solution $u \in X$ de

$$F(u) \leq F(v) \quad \forall v \in X$$

est solution de

$$(F'(u), v-u) \geq 0 \quad \forall v \in X$$

ou de

$$(F'(v), v-u) \geq 0 \quad \forall v \in X$$

et réciproquement.

2°) Si $F = F_1 + F_2$ avec F_1 différentiable alors toute solution $u \in X$ de

$$F(u) \leq F(v) \quad \forall v \in X$$

est solution de

$$(-F'_1(u), v-u) \leq F_2(v) - F_2(u) \quad \forall v \in X$$

ou de

$$(-F'_1(v), v-u) \leq F_2(v) - F_2(u) \quad \forall v \in X$$

et réciproquement.

3°) Si F n'est pas différentiable alors toute solution $u \in X$ de

$$F(u) \leq F(v) \quad \forall v \in X$$

est solution de

$$(-Au, v-u) \leq G(v) - G(u) \quad \forall v \in X$$

ou de

$$(-Av, v-u) \leq G(v) - G(u) \quad \forall v \in X$$

et réciproquement.

Les opérateurs A et G étant ceux du théorème 3.1.

Dans les paragraphes qui suivent on se propose de donner quelques méthodes permettant de résoudre l'inéquation

$$(Au, v-u) \geq (f, v-u) \quad \forall v \in X \text{ pour } f \text{ donné dans } V'.$$

I-4. Une méthode d'approximation par régularisation elliptique.

Soit V un Banach uniformément convexe ; K est un convexe fermé de V contenant l'origine. On fait les hypothèses suivantes :

H.4) B est un opérateur de K dans V' monotone hémicontinu tel que

$$(Bu - Bv, u - v) \geq (\varphi(\|u\|) - \varphi(\|v\|)) (\|u\| - \|v\|)$$

avec φ strictement croissante de $\mathbb{R}_+$ dans $\mathbb{R}$ tel que $\lim_{r \rightarrow \infty} \varphi(r) = +\infty$.

H.5) Soit $A : K \longrightarrow V'$ monotone hémicontinu. Soit X l'ensemble convexe fermé des solutions de $u \in K$

$$(4.1) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in K ; \text{ pour } f \text{ donné dans } V'.$$

H.6) On suppose $X \neq \emptyset$.

Pour g donné dans V' , d'après le théorème 1.1, $\exists u_0 \in X$ tel que

$$(4.2) \quad (Bu_0, v - u_0) \geq (g, v - u_0) \quad \forall v \in X.$$

un tel u_0 est unique en raison de H.4).

On va donner un théorème d'approximation de u_0 , adaptation au cas non linéaire de la méthode de [20].

Théorème 4.1

On suppose H.4) et H.5) alors

1°) $\forall \varepsilon > 0$, $\exists u_\varepsilon \in K$ unique solution de

$$(4.3) \quad (Au_\varepsilon, v - u_\varepsilon) + \varepsilon (Bu_\varepsilon, v - u_\varepsilon) \geq (f + \varepsilon g, v - u_\varepsilon) \\ \forall v \in K, \text{ pour } f \text{ et } g \text{ donnés dans } V'.$$

2°) Si de plus on a H.6) alors

$u_\varepsilon \longrightarrow u_0$ dans V fort quand $\varepsilon \longrightarrow 0$, où u_0 est une solution de (4.1) vérifiant (4.2).

Démonstration

1°) L'existence et l'unicité de u_ε résulte du théorème 1.1.

2°) a) Montrons que (u_ε) est borné dans V

En effet, il résulte de (4.1) et (4.3) que

$$(Au_\varepsilon - Au_0, u_0 - u_\varepsilon) + \varepsilon (Bu_\varepsilon, u_0 - u_\varepsilon) \geq \varepsilon (g, u_0 - u_\varepsilon)$$

Donc $(g - Bu_\varepsilon, u_0 - u_\varepsilon) \leq 0$.

Il en résulte que

$$(Bu_\varepsilon - Bu_0, u_\varepsilon - u_0) \leq (g - Bu_0, u_\varepsilon - u_0)$$

Comme

$$(\varphi(\|u_\varepsilon\|) - \varphi(\|u_0\|)) (\|u_\varepsilon\| - \|u_0\|) \leq (Bu_\varepsilon - Bu_0, u_\varepsilon - u_0),$$

et donc

$$(4.4) \quad \left(\varphi(\|u_\varepsilon\|) - \varphi(\|u_0\|) \right) \left(1 - \frac{\|u_0\|}{\|u_\varepsilon\|} \right) \leq \left(g - Bu_0, \frac{u_\varepsilon - u_0}{\|u_\varepsilon\|} \right)$$

Il résulte de (4.4) que $\|u_\varepsilon\| \leq cte$.

Donc il existe un ultrafiltre $\mathcal{U}$, tel que $\lim_{\mathcal{U}} u_\varepsilon = \xi$ dans V faible.

b) Montrons que $\xi \in X$

La relation (4.3) est équivalente d'après la remarque 1.1 à

$$(4.5) \quad (Av + \varepsilon Bv, v - u_\varepsilon) \geq (f + \varepsilon g, v - u_\varepsilon) \quad \forall v \in K$$

En passant à la limite suivant $\mathcal{U}$ on a :

$$(Av, v - \xi) \geq (f, v - \xi) \quad \forall v \in K$$

ce qui est équivalent à

$$(A\xi, v - \xi) \geq (f, v - \xi) \quad \forall v \in K, \text{ d'où } \xi \in X.$$

c) $\|u_\varepsilon\| \longrightarrow \|\xi\|$ quand $\varepsilon \longrightarrow 0$

On fait $v = \xi$ dans (4.3) et dans (4.1) on fait

$v = u_\varepsilon$ et $u = \xi$. Par addition il vient :

$$(Au_\varepsilon - A\xi, \xi - u_\varepsilon) + \varepsilon (Bu_\varepsilon, \xi - u_\varepsilon) \geq \varepsilon (g, \xi - u_\varepsilon)$$

D'après la monotonie de A , nous avons :

$$(4.6) \quad \varepsilon (g - Bu_\varepsilon, \xi - u_\varepsilon) \leq (Au_\varepsilon - A\xi, \xi - u_\varepsilon) \leq 0.$$

Posons $u = u_\varepsilon$ et $v = \xi$ dans H.4), alors

$$(\varphi(\|u_\varepsilon\|) - \varphi(\|\xi\|)) (\|u_\varepsilon\| - \|\xi\|) \leq (Bu_\varepsilon - B\xi, u_\varepsilon - \xi)$$

Posons $v = u_\varepsilon$ dans (4.1) ; par addition avec (4.5) où on a fait $v = u$ on a :

$$(4.7) \quad (Bu, u - u_\varepsilon) \geq (g, u - u_\varepsilon) \quad \forall u \in X.$$

En passant à la limite suivant $\mathcal{U}$:

$$(Bu, u - \xi) \geq (g, u - \xi) \quad \forall u \in X. \text{ Ce qui équivaut à :}$$

$$(B\xi, u - \xi) \geq (g, u - \xi) \quad \forall u \in X.$$

Donc ξ est solution de (4.2) et d'après l'unicité on a $\xi = u_0$.

Par ailleurs de (4.6) on déduit

$$(Bu_\varepsilon, u_\varepsilon - \xi) \leq (g, u_\varepsilon - \xi)$$

ce qui donne

$$0 \leq (\varphi(\|u_\varepsilon\|) - \varphi(\|\xi\|)) (\|u_\varepsilon\| - \|\xi\|) \leq (g - B\xi, u_\varepsilon - \xi) \longrightarrow 0 \text{ suivant } \mathcal{U}.$$

Il résulte de la stricte monotonie de φ que $\|u_\varepsilon\| \longrightarrow \|\xi\|$.

Donc $u_\varepsilon \longrightarrow u_0$ suivant $\mathcal{U}$ dans V fort.

D'où le théorème.

Théorème 4.2

On suppose H.4) et H.5) alors

X non vide $\iff$ l'ensemble $\{u_\varepsilon\}$ borné

Démonstration.

1°) Si $X \neq \emptyset$; alors $\|u_\xi\| \leq \text{cte}$ d'après le théorème 4.1 2°), a).

2°) Si $\|u_\xi\| \leq \text{cte}$; alors suivant $\mathcal{U}$ on a

$u_\xi \rightarrow \xi$ et on montre que $\xi \in X$.

I - 5. Une méthode d'approximation par discrétisation.

Soit $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ un paramètre destiné à tendre vers zéro ; $h > 0$. On se propose dans ce paragraphe d'approcher la solution de l'inéquation

$$(Au, v-u) \geq (f, v-u) \quad \forall v \in X$$

pour f donné dans V' et $A : X \rightarrow V'$ par la solution de l'inéquation plus simple à résoudre

$$(A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

où X_h est un convexe fermé d'un espace de dimension finie V_h et A_h un opérateur de X_h dans V'_h .

Plus précisément soit V un Banach sur $\mathbb{R}$, uniformément convexe de norme $\|\cdot\|$. On se donne un convexe X fermé de V contenant l'origine.

Soit V_h un espace de dimension finie (associé au paramètre h) de dual V'_h ; muni de la norme $\|\cdot\|_h$. On se donne un prolongement $p_h \in \mathcal{L}(V_h, V)$ injectif de V_h dans V . On note par $\hat{p}_h$ sa transposée cf [1].

Posons $\|u_h\|_h = \|p_h u_h\| \quad \forall u_h \in V_h$

$$A_h = \hat{p}_h A p_h \quad ; \quad X_h = p_h^{-1}(X)$$

Nous avons alors le schéma suivant :

$$\begin{array}{ccc} V & \xrightarrow{A} & V' \\ p_h \uparrow & & \downarrow \hat{p}_h \\ V_h & \xrightarrow{A_h} & V'_h \end{array}$$

Nous notons par $(\cdot, \cdot)_h$ le produit scalaire qui met V_h et V'_h en dualité. Pour $f_h \in V'_h$ on pose :

$$\|f_h\|_h^* = \sup_{v_h \in V_h} \frac{|(f_h, v_h)_h|}{\|v_h\|_h}$$

Posons d'après [1] la

Définition 5.1 : Soient $f_h \in V'_h$ et $f \in V'$. On dit que f_h converge discrètement vers f si $\lim_{h \rightarrow 0} \|\hat{p}_h f - f_h\|_h^* = 0$

On fait alors les hypothèses suivantes :

H'2) A est un opérateur monotone hémicontinu de X dans V' tel que

$$\lim_{\substack{\|u\| \rightarrow \infty \\ u \in X}} \frac{(Au, u)}{\|u\|} = +\infty$$

Si pour une suite x_1 on a : $(Ax_1 - Au, x_1 - u) \longrightarrow 0$ alors $\|x_1\| \longrightarrow \|u\|$.

Il est clair que $H'2) \implies H2)$.

H.7) A transforme tout borné de V en un borné de V'.

Nous avons alors le

Théorème 5.1

On suppose H'2) et H 7). De plus on suppose

H 8) $\forall x \in X, \exists u_h \in X_h$ tel que $p_h u_h \longrightarrow x$ dans V fort.

Alors

1°) $\forall f_h \in V'_h \quad \exists u_h \in X_h$ unique tel que

$$(5.1) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

2°) De plus si f_h converge discrètement vers f on a :

$$p_h u_h \longrightarrow u \text{ dans V fort quand } h \longrightarrow 0.$$

où u est l'unique solution dans X de l'inéquation

$$(5.2) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X.$$

Démonstration.

1°) Le théorème 1.1 d'existence et d'unicité s'applique directement à (5.1).

2°) a) Montrons que $\|p_h u_h\|$ est borné.

Nous avons :

$$\frac{(A_h u_h, u_h)_h}{\|u_h\|_h} = \frac{(A p_h u_h, p_h u_h)}{\|p_h u_h\|} \leq \frac{(f_h, p_h u_h)}{\|p_h u_h\|} \leq \|f_h\|_h^*$$

Si f_h converge discrètement vers f alors $\|f_h\|_h^*$ est borné. Supposons maintenant

que $\|p_h u_h\|$ ne soit pas borné, alors il existerait une suite extraite encore

notée $p_h u_h$ telle que $\|p_h u_h\| \longrightarrow \infty$ et alors $\frac{(A_h u_h, u_h)_h}{\|u_h\|_h} \longrightarrow \infty$; ce qui

est impossible. Donc $\|p_h u_h\| \leq \text{cte.}$

b) Donc suivant un ultrafiltre $p_h u_h \longrightarrow \xi$ dans X faible.

D'après H.7) : $Ap_h u_h \longrightarrow \eta$ dans V' faible. Puisque $\xi \in X$, alors $\exists \xi_h \in X_h$ tel que $p_h \xi_h \longrightarrow \xi$ dans V fort.

Nous avons alors

$$(5.3) \quad (Ap_h u_h, p_h u_h) = (A_h u_h, u_h)_h \leq (f_h, u_h - v_h)_h + (A_h u_h, v_h)_h$$

Faisons $v_h = \xi_h$ dans (5.3) il vient alors

$$\limsup (Ap_h u_h, p_h u_h) \leq \limsup (f_h, u_h - \xi_h)_h + \limsup (Ap_h u_h, p_h \xi_h)$$

comme $\limsup (f_h, u_h - \xi_h) \longrightarrow 0$ il résulte que

$\limsup (A p_h u_h, p_h \xi_h) \leq (\eta, \xi)$. Or si A est monotone hemicontinu et si un filtre x_i converge vers x dans V faible, $Ax_i \rightarrow y$ dans V' faible et $\limsup (Ax_i, x_i) \leq (y, x)$ alors $Ax = y$. On en déduit alors que $A\xi = \eta$.

c) Montrons la convergence forte.

Posons

$$\alpha_h = (Ap_h u_h - A\xi, p_h u_h - \xi). \text{ Nous avons alors}$$

$$\alpha_h = (Ap_h u_h, p_h u_h) - (Ap_h u_h, \xi) + (A\xi, \xi) - (A\xi, p_h u_h)$$

et $\limsup \alpha_h \leq 0$. D'où

$$(Ap_h u_h - A\xi, p_h u_h - \xi) \longrightarrow 0 \text{ quand } h \longrightarrow 0.$$

Il résulte de H'2) que $\|p_h u_h\| \longrightarrow \|\xi\|$

Comme V est uniformément convexe et $p_h u_h \longrightarrow \xi$ dans V faible ; alors

$p_h u_h \longrightarrow \xi$ dans V fort.

Soit $v \in X$. D'après l'hypothèse H.8), $\exists v_h \in X_h$ tel que $p_h v_h \longrightarrow v$ dans V fort.

En passant à la limite dans

$$(A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \text{ on a :}$$

$$(A\xi, v - \xi) \geq (f, v - \xi). \text{ D'après l'unicité on a : } \xi = u.$$

Donc $p_h u_h \longrightarrow u$ dans V fort.

Remarque 5.1

Si $X = V$ l'hypothèse H.8) est réalisée si

$$\lim_{h \rightarrow 0} \|p_h r_h u - u\| = 0 \quad \forall u \in V \text{ où } r_h \in \mathcal{L}(V, V_h)$$

appelée restriction de V dans V_h .

Remarque 5.2

Si $X \neq V$ l'hypothèse H.8) est réalisée, si l'approximation

$\{V, V_h, p_h, r_h\}$ est stable et consistante (cf [1]) et si de plus

$$p_h(X_h) \subset p_h(X_{h'}) \text{ pour } h \leq h'$$

et $\bigcup_h p_h(X_h)$ est dense dans X .

Soit V un Banach réflexif de norme $\| \cdot \|$. Soit $h > 0$ un paramètre (destiné à tendre vers 0). Nous introduisons comme précédemment le quadruplet $\{V, V_h, p_h, r_h\}$ où V_h est un espace de dimension finie de norme $\| \cdot \|_h$ $p_h \in \mathcal{L}(V_h, V)$ appelé prolongement de V_h dans V et $r_h \in \mathcal{L}(V, V_h)$ appelé restriction de V dans V_h .

On désigne par $\hat{p}_h$ et $\hat{r}_h$ les transposées de p_h et r_h respectivement. Les applications p_h et r_h sont dites stables si on a :

$$\|p_h u_h\| \leq c_1 \|u_h\|_h \text{ et } \|r_h u\|_h \leq c_2 \|u\|$$

où c_1 et c_2 sont des constantes indépendantes de h . On sait cf [1] que les familles $\{V, V_h, p_h, r_h\}$ et $\{V', V'_h, \hat{p}_h, \hat{r}_h\}$ sont stables en même temps.

Nous rappelons alors par commodité pour le lecteur les définitions suivantes :

Définition 5.2

La famille $\{V, V_h, p_h, r_h\}$ est dite consistante si

$$\lim_{h \rightarrow 0} \|p_h r_h u - u\| = 0 \quad \forall u \in V$$

Définition 5.3

$u_h \in V_h$ converge discrètement vers $u \in V$ si

$$\lim_{h \rightarrow 0} \|r_h u - u_h\|_h = 0$$

On a une définition analogue pour la convergence discrète de $f_h \in V'_h$ vers $f \in V'$. Dans ce cas on montre facilement que $\|f_h\|_h^*$ est borné.

Définition 5.4

On dit que l'opérateur A_h de V_h dans V'_h est consistant à l'opérateur $A : V \rightarrow V'$ si

$$\lim_{h \rightarrow 0} \|\hat{p}_h A u - A_h r_h u\|_h^* = 0 \quad \forall u \in V$$

Dans toute la suite de ce paragraphe nous allons nous occuper plus particulièrement et uniquement pour simplifier, de l'approximation de $u \in V$ solution de l'équation

$$(5.4) \quad Au = f \text{ pour } f \text{ donné dans } V', A \text{ monotone hémicontinu.}$$

On suppose maintenant que (5.4) admet une solution unique $u \in V$. Considérons alors l'équation

$$(5.5) \quad A_h u_h = f_h \text{ pour } f_h \text{ donné dans } V'_h$$

Sous certaines conditions, nous allons montrer que (5.5) admet une solution unique $u_h \in V_h$; que $p_h u_h \rightarrow u$ dans V fort quand $h \rightarrow 0$ et on a des estimations de l'erreur sur la discrétisation.

On fait les hypothèses suivantes :

- K1) La famille $\{V, V_h, p_h, r_h\}$ est stable et consistante.
- K2) A_h est hémicontinu de V_h dans V'_h et vérifie
 $(A_h u_h - A_h v_h, u_h - v_h)_h \geq C \|u_h - v_h\|_h^2 \quad C > 0$
- K3) A_h est consistant à l'opérateur A .

Posons les notations suivantes :

$$\|p_h u_h\| \leq C' \|u_h\|_h$$

$$\alpha(h) = \|\hat{f}_h A u - A_h r_h u\|_h^*$$

$$\beta(h) = \|p_h r_h u - u\|$$

$$\text{On suppose : } \gamma(h) = \|\hat{f}_h - \hat{f}_h \hat{f}\|_h^* \longrightarrow 0 \quad \text{quand } h \longrightarrow 0$$

Nous avons alors le

Théorème 5.2

On suppose K1), K2), K3) alors

- 1) Il existe $u_h \in V_h$ unique solution de (5.5)
- 2) $p_h u_h \rightarrow u$ dans V fort quand $h \rightarrow 0$ u étant la solution de (5.4).
- 3) On a : $\|p_h u_h - u\| \leq \frac{C'}{C} (\alpha(h) + \gamma(h)) + \beta(h)$

Démonstration

- 1) L'existence et l'unicité de $u_h \in V_h$ est évidente.
- 2) Montrons que u_h converge vers u discrètement.

Nous avons :

$$C \|u_h - r_h u\|^2 \leq (A_h u_h - A_h r_h u, u_h - r_h u)_h$$

Ce qui donne :

$$\|u_h - r_h u\|_h \leq \frac{1}{c} \|A_h u_h - A_h r_h u\|_h^* \leq \frac{1}{c} (\alpha(h) + \gamma(h)) \longrightarrow 0$$

quand $h \longrightarrow 0$ d'après K.3) et la convergence discrète de f_h .

3) $p_h u_h \longrightarrow u$ dans V fort. En effet :

$$\|p_h u_h - u\| \leq \|p_h u_h - p_h r_h u\| + \|p_h r_h u - u\| \leq \frac{c'}{c} (\alpha(h) + \gamma(h)) + \beta(h) \longrightarrow 0$$

quand $h \longrightarrow 0$ d'après K.1) et 2). D'où le théorème.

Remarque 5.3

Dans le théorème précédent nous pouvons munir V_h de la norme $\|u_h\|_h = \|p_h u_h\|$ si p_h est injective, et poser

$$A_h = \hat{p}_h A p_h ; f_h = \hat{p}_h f, \text{ ce qui donne :}$$

$$\|u_h - r_h u\| \leq \frac{1}{c} \alpha(h) \quad \text{et} \quad \|p_h u_h - u\| \leq \frac{1}{c} \alpha(h) + \beta(h)$$

$$\text{avec } \alpha(h) \leq \|Au - A p_h r_h u\|.$$

Plus précisément si A est hémicontinu, lipschitzien et vérifie $(Au - Av, u-v) \geq c \|u-v\|^2$ $c > 0$; de plus si on suppose K.1) et $A_h = \hat{p}_h A p_h$, $f_h = \hat{p}_h f$, $\|u_h\|_h = \|p_h u_h\|$ alors $A_h u_h = f_h$ admet une solution unique $u_h \in V_h$ et on a la majoration à priori de l'erreur :

$$\|p_h u_h - u\| \leq K \beta(h).$$

Dans [2.] J.P. AUBIN donne des majorations de $\beta(h)$ en fonction de h dans le cas où $V = H^m(\Omega) = \{u | u \in L^2(\Omega), D_p u \in L^2(\Omega), |p| \leq m\}$

Nous donnons maintenant le "discrétisé" du théorème 1.2. On suppose : K.4) V est un Banach uniformément convexe. A est un opérateur monotone hémicontinu de B_r dans V' ; vérifiant $(Au, u) \geq 0 \quad \forall u \in S_r$.

De plus si pour un filtre x_i on a $(Ax_i - Au, x_i - u) \longrightarrow 0$ alors $\|x_i\| \longrightarrow \|u\|$.

K.5) La famille $\{V, V_h, p_h, r_h\}$ est consistante

On pose :

$$\|u_h\|_h = \|p_h u_h\| ; B_h = p_h^{-1}(B_r) ; S_r = p_h^{-1}(S_r)$$

$$A_h = \hat{p}_h A p_h = B_h \longrightarrow V'_h.$$

Théorème 5.3

On suppose K.4) et K.5) alors

1°) Il existe $u_h \in B_h$ unique tel que $A_h u_h = 0$

2°) Si A est un opérateur borné (transforme un borné de V en un borné de V') alors $p_h u_h \longrightarrow u$ dans V fort où u est solution de $Au = 0$.

Démonstration.

1°) Cette assertion se déduit directement du théorème 1.2

2°) Montrons la convergence forte. Comme $u_h \in B_h$ alors $\|u_h\|_h = \|p_h u_h\| \leq r$

Donc suivant un ultra filtre $p_h u_h \longrightarrow \xi$ dans V faible et

$A p_h u_h \longrightarrow \eta$ dans V' faible. Or

$$0 = (A_h u_h, r_h v)_h = (A p_h u_h, p_h r_h v) \longrightarrow (\eta, v) \quad \forall v.$$

Donc $\eta = 0$. Comme $\limsup (A p_h u_h, p_h u_h) = 0$, il résulte que $A\xi = 0$.

Donc $\xi = u$ d'après l'unicité de la solution u de l'équation $Au = 0$. Par ailleurs nous avons :

$\lim_{h \rightarrow 0} (A p_h u_h - Au, p_h u_h - u) = 0 \implies \|p_h u_h\| \longrightarrow \|u\|$; et d'après l'uniforme convexité de V on a :

$p_h u_h \longrightarrow u$ dans V fort. D'où le théorème.

Remarque 5.4

Si au lieu de l'hypothèse K2) on suppose

$$K'2) (A_h u_h - A_h v_h, u_h - v_h)_h \geq (\varphi(\|u_h\|_h) - \varphi(\|v_h\|_h)) (\|u_h\|_h - \|v_h\|_h)$$

où φ est strictement croissante de $\mathbb{R}^+$ dans $\mathbb{R}$ et tel que $\lim_{r \rightarrow \infty} \varphi(r) = +\infty$

1) on peut se ramener à l'hypothèse K2) par régularisation elliptique (cf. théorème 4.1) : L'opérateur $B_h = A_h + \lambda(h).Id$ avec $\lambda(h) \rightarrow 0$ quand $h \rightarrow 0$, vérifie alors K2.

2) si $A_h = \hat{r}_h A p_h$ et A vérifie K'2) donc A_h aussi, alors $B = A + \lambda Id$ et $B_h = A_h + \lambda Id$ vérifient K2) et la solution $u_{h,\lambda}$, de l'équation $B_h u_{h,\lambda} = f_h$, est telle que $p_h u_{h,\lambda} \rightarrow u_\lambda$, solution de $B u_\lambda = f$. Le théorème 4.1 dit que $u_\lambda \rightarrow u$ dans V fort quand $\lambda \rightarrow 0$, u étant la solution de $Au = f$.

3) le paramètre λ peut aussi être choisi suffisamment grand de manière à minimiser l'erreur de discrétisation $\|p_h u_h - u\|$, puisque $\lambda = c$, constante intervenant dans l'hypothèse K2).

I - 6. Méthode de pénalisation des contraintes.

Soit V un Banach réflexif de dual V' . Soit A un opérateur non nécessairement linéaire de V dans V' et X un convexe fermé de V . On cherche $u \in X$ solution de l'inégalité :

$$(6.1) \quad (A u, v-u) \geq (f, v-u) \quad \forall v \in X, \text{ pour } f \text{ donné dans } V'.$$

Soit $\varepsilon > 0$ un paramètre destiné à tendre vers 0. On se propose maintenant d'approcher $u \in X$ solution de (6.1) par $u_\varepsilon \in V$ solution d'une équation de la forme :

$$(6.2) \quad A u_\varepsilon + \frac{1}{\varepsilon} B u_\varepsilon = f$$

Dans ce nouveau problème les contraintes ont disparues. Ce procédé porte le nom de "pénalisation des contraintes" cf. [16] et l'idée en remonte à Courant [12].

Plus précisément on fait les hypothèses suivantes :

i) Le convexe X est de la forme

$$X = \{ v \in V \mid B v = 0 \}$$

où B est un opérateur monotone et hémicontinu.

ii) A est un opérateur hémicontinu tel que

$$(A u - A v, u-v) \geq (\varphi(\|u\|) - \varphi(\|v\|)) (\|u\| - \|v\|) \quad \forall u, v \in V$$

où φ est une application de $\mathbb{R}_+$ dans $\mathbb{R}$ strictement croissante telle que $\lim_{r \rightarrow \infty} \varphi(r) = +\infty$

iii) V est un espace uniformément convexe.

Nous avons alors le

Théorème 6.1

1°) Il existe $u \in X$ unique solution de (6.1).

2°) Pour tout $\varepsilon > 0$ $\exists u_\varepsilon \in V$ unique solution de (6.2).

3°) $u_\varepsilon \longrightarrow u$ dans V fort quand $\varepsilon \rightarrow 0$.

Démonstration

Pour avoir les deux premières assertions, il suffit d'appliquer le théorème 1.1.

3°) Soit K l'ensemble convexe des solutions de (6.1) (ici réduit à un point). Toute solution de (6.2) est solution de

$$(6.3) \quad \varepsilon A u_\varepsilon + B u_\varepsilon = \varepsilon f$$

ou de

$$(6.4) \quad (B u_\varepsilon + \varepsilon A u_\varepsilon, v - u_\varepsilon) \geq \varepsilon (f, v - u_\varepsilon) \quad \forall v \in V$$

et réciproquement. On applique alors le théorème 4.1 qui dit que $u_\varepsilon \longrightarrow u$ dans V fort quand $\varepsilon \rightarrow 0$ tel que $u \in X$ c'est-à-dire $B u = 0$ et vérifiant $(A u, v - u) \geq (f, v - u) \quad \forall v \in X$. D'où le théorème

Remarque 6.1

La démonstration du théorème 6.1 déduite de celle du théorème 4.1 montre l'analogie entre la régularisation elliptique et la pénalisation des contraintes.

On trouvera dans [21] une démonstration directe du théorème 6.1.

Corollaire 6.1

Soit X un convexe fermé d'un espace de Hilbert H de norme $\| \cdot \|$ et P l'opérateur de projection de H sur le convexe X . On se donne $A : H \rightarrow H'$ vérifiant ii) alors

1°) Il existe $u \in X$ unique solution de

$$(6.5) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

2°) Pour tout $\varepsilon > 0$ $\exists u_\varepsilon \in H$ unique solution de

$$(6.6) \quad A u_\varepsilon + \frac{1}{\varepsilon} (u_\varepsilon - P u_\varepsilon) = 0$$

3°) $u_\varepsilon \longrightarrow u$ dans H fort quand $\varepsilon \rightarrow 0$.

Démonstration

Posons $B = I - P$. C'est un opérateur monotone hémicontinu car $\| P u - P v \| \leq \| u - v \|$. De même, nous avons :

$$X = \{ u \mid u \in H ; P u = u \} = \{ u \in H \mid B u = 0 \}$$

On applique alors le théorème 6.1.

Remarque 6.2

Le théorème (6.1) est encore valable dans le cas où

$$X = \{ v \mid B v = 0 \} \quad \text{avec } B \text{ monotone}$$

hémicontinu et multivoque.

On démontre dans [21] le :

Corollaire 6.2

On suppose $X = \{ v \in V \mid g(v) = 0 \}$

avec $g \geq 0$ convexe s.c.i et dérivable au sens de Gateaux.

Soit $F : X \rightarrow \mathbb{R}$ une fonctionnelle strictement convexe s.c.i non identique à $+\infty$ dérivable au sens de Gateaux et telle que

$$\lim_{\substack{\|u\| \rightarrow \infty \\ u \in X}} F(u) = +\infty. \text{ Alors}$$

1°) Il existe $u \in X$ unique solution de

$$(6.7) \quad F(u) \leq F(v) \quad \forall v \in X$$

2°) Pour tout $\varepsilon > 0$ $\exists u_\varepsilon \in V$ unique solution de

$$(6.8) \quad F(u_\varepsilon) + \frac{1}{\varepsilon} g(u_\varepsilon) \leq F(v) + \frac{1}{\varepsilon} g(v) \quad \forall v \in V$$

3°) $u_\varepsilon \rightarrow u$ dans V fort quand $\varepsilon \rightarrow 0$

Dans le cas où F et g ne sont pas différentiables au sens de Gateaux on a le

Théorème 6.2

On suppose $X = \{ v \in V \mid g(v) = 0 \}$ avec $g \geq 0$ convexe et s.c.i

Soit $F : X \rightarrow \mathbb{R}$ une fonctionnelle strictement convexe s.c.i telle que

$$\lim_{\substack{\|u\| \rightarrow \infty \\ u \in X}} F(u) = +\infty.$$

Si de plus on suppose la condition : α) $u_\varepsilon \rightarrow w$ dans V faible et $F(u_\varepsilon) \rightarrow F(w)$ entraîne que $u_\varepsilon \rightarrow w$ dans V fort; alors

1°) Il existe $u \in X$ unique solution de (6.7)

2°) Pour tout $\varepsilon > 0$ $\exists u_\varepsilon \in V$ unique solution de (6.8)

3°) $u_\varepsilon \rightarrow u$ dans V fort quand $\varepsilon \rightarrow 0$.

Démonstration.

Les deux premières assertions découlent du théorème 2.2

3°) Soit $v_0 \in X$ on a :

$$(6.9) \quad F(u_\varepsilon) + \frac{1}{\varepsilon} g(u_\varepsilon) \leq F(v_0)$$

Donc $\|u_\varepsilon\| \leq c^{\text{te}}$ sinon le premier membre de (6.9) tendrait vers $+\infty$ ce qui est impossible.

Donc suivant un ultrafiltre $\mathcal{U}$ on a $u_\varepsilon \rightarrow \xi$ dans X . Par ailleurs, nous avons

$$(6.10) \quad F(\xi) \leq \liminf_{\mathcal{U}} F(u_\varepsilon) \leq F(v_0) \quad \forall v_0 \in X$$

Comme

$$\varepsilon F(u_\varepsilon) + g(u_\varepsilon) \leq \varepsilon F(v) + g(v) \quad \forall v \in V$$

on a à la limite :

$$g(\xi) \leq \liminf_{\mathcal{U}} g(u_\varepsilon) \leq g(v) \quad \forall v \in V$$

Donc $g(\xi) = 0$ et $\xi \in X$. Comme la solution est unique, on a $\xi = u$ et donc $u_\varepsilon \rightarrow u$ dans V faible. Il suffit de montrer maintenant que $F(u_\varepsilon) \rightarrow F(u)$ pour avoir la convergence forte de u_ε vers u . Nous procédons successivement

$$(6.11) \quad F(u) \leq \liminf_{\mathcal{U}} F(u_\varepsilon)$$

$$(6.12) \quad \limsup_{\mathcal{U}} F(u_\varepsilon) \leq F(v_0) \quad \forall v_0 \in X$$

Donc $\lim_{\mathcal{U}} F(u_\varepsilon) = F(u)$ Donc $u_\varepsilon \rightarrow u$ dans V fort.

Remarque 6.3

La condition $\alpha)$ du théorème précédent se trouve toujours satisfaite si l'espace V est uniformément convexe et la fonctionnelle F est de la forme

$$F(u) = \|u\|^p + F_0(u) \quad \text{avec } F_0 \text{ convexe s.c.i.}$$

En effet, nous avons

$$\liminf \|u_\varepsilon\|^p \geq \|u\|^p,$$

$$\limsup \|u_\varepsilon\|^p \leq \limsup (\|u_\varepsilon\|^p + F_0(u_\varepsilon)) - \liminf F_0(u_\varepsilon)$$

$$\limsup \|u_\varepsilon\|^p \leq \|u\|^p + F_0(u) - F_0(u) = \|u\|^p.$$

Donc $\|u_\varepsilon\|^p \rightarrow \|u\|^p$ quand $\varepsilon \rightarrow 0$.

I - 7. Méthode d'éclatement et pénalisation partielle des contraintes.

Soit V un Banach uniformément convexe de dual V' et A un opérateur non nécessairement linéaire de V dans V' . On se propose de résoudre l'inégalité

$$(7.1) \quad (A u, v-u) \geq (f, v-u) \quad \forall v \in X \text{ pour } f \text{ donné dans } V' \text{ où } X \text{ est un convexe fermé de } V \text{ de la forme :}$$

$$(7.2) \quad X = \bigcap_{i=1}^q X_i \text{ avec } X_i \text{ convexe fermé de } V; i = 1, \dots, q.$$

Dans un grand nombre d'applications, il arrive que le convexe X soit un ensemble avec beaucoup de "contraintes". Il est bon alors d'utiliser la pénalisation uniquement pour certaines contraintes. Plus précisément :

Supposons que le convexe X soit de la forme :

$$(7.3) \quad X = X_1 \cap X_2 \text{ avec } X_1 = \bigcap_{i=1}^r X_i, X_2 = \bigcap_{j=1}^s X_j \quad r + s = q$$

où les X_i sont des convexes sur lesquels on sait projeter et les X_j sont des convexes sur lesquels la projection numérique n'est pas immédiate.

Nous allons maintenant appliquer une méthode due à [18] et ramener la résolution de (7.1) à celles de q problèmes de la forme :

$$(7.4)_i \quad (A_i u, v-u) \geq (f_i, v-u) \quad \forall v \in X_i \quad i = 1, \dots, q$$

On a alors les deux possibilités :

- 1) Si la projection sur X_i est immédiate pour $i = i_0$; on applique les méthodes de [22], pour résoudre par itérations l'inégalité $(7.4)_{i_0}$
- 2) Si pour $i = j_0$ la projection sur le convexe X_{j_0} est difficile on peut alors par "pénalisation" ramener l'inégalité $(7.4)_{j_0}$ à l'équation sans contraintes de la forme :

$$(7.5)_{j_0} \quad A_{j_0} u + \frac{1}{\epsilon} B_{j_0} u = f_{j_0} \quad (\text{avec } \epsilon > 0 \text{ destiné à tendre vers } 0)$$

que l'on sait résoudre par une méthode itérative cf. [6].

Evidemment il y a une relation entre la solution du problème (7.1) et les solutions des problèmes $(7.4)_i$ $i = 1, \dots, q$ et éventuellement $(7.5)_{j_0}$..

Plus précisément, soient V_i ($i = 1, \dots, q$) des Banach uniformément convexes de normes $\| \cdot \|_i$. On pose

$$V = \bigcap_{i=1}^q V_i. \text{ L'espace } V \text{ est muni de la norme } \| \cdot \| \text{ équivalente à } \sum_{i=1}^q \| \cdot \|_i.$$

On se donne un Hilbert H muni de la norme $\| \cdot \|$, avec $V \subset V_i \subset H \subset V'_i \subset V'$; les injections étant continues, $(\cdot, \cdot)$ désigne le produit scalaire dans H , dans la dualité entre V_i, V'_i et V, V' .

Soit X un convexe fermé de V . On suppose

$$(7.6) \quad X = \bigcap_{i=1}^q X_i ; \text{ avec } X_i \text{ convexe fermé de } V_i, \text{ contenant l'origine.}$$

Soient $A : X \rightarrow V'$ et $A_i : X_i \rightarrow V'_i \quad i = 1, \dots, q$

tels que

$$(7.7) \quad A = \sum_{i=1}^q A_i \quad \text{et} \quad f = \sum_{i=1}^q f_i$$

$$(7.8) \quad (A_i u - A_i v, u-v) \geq (\varphi_i(\|u\|_i) - \varphi_i(\|v\|_i)) (\|u\|_i - \|v\|_i)$$

avec φ_i strictement croissante de $\mathbb{R}_+$ dans $\mathbb{R}$ et tel que $\lim_{r \rightarrow \infty} \varphi_i(r) = +\infty$.
On se donne $u^0 \in X_q$. Ayant déterminé $u^{n+\frac{i-1}{q}} \in X_{i-1}$ on obtient

$u^{n+\frac{i}{q}}$ comme solution dans X_i de l'inéquation (ε_0 donné positif) $k > 0$

$$(7.9)_i \quad \left((A_i + \frac{\varepsilon_0}{k}) u^{n+\frac{i}{q}}, v - u^{n+\frac{i}{q}} \right) \geq (f_i + \frac{\varepsilon}{k} u^{n+\frac{i-1}{q}}, v - u^{n+\frac{i}{q}}) \quad \forall v \in X_i$$

Nous avons alors le

Théorème 7.1

On suppose (7.6), (7.7), (7.8). Soit $k = \frac{1}{N} \cdot T$; T fixé et N entier.
On se donne deux paramètres ε_0 , tels que $\varepsilon_0 \rightarrow 0, k \rightarrow 0$
avec $\frac{k}{\varepsilon_0} \rightarrow 0$. Alors

1°) Quel que soit $i \quad (i = 1, \dots, q)$

$$(7.10) \quad \frac{1}{N} \sum_{n=0}^{N-1} u^{n+\frac{i}{q}} \longrightarrow u \quad \text{dans } V_i \text{ fort, } u \text{ étant la solution de}$$

$$(7.11) \quad (A u, v-u) \geq (f, v-u) \quad \forall v \in X$$

$$\text{Posons } B_i = A_i + \frac{\varepsilon_0}{k} \text{Id}$$

2°) On suppose maintenant que $V_i \quad i = 1, \dots, q$ sont des espaces de Hilbert et que $B_i : V_i \rightarrow V'_i$ vérifient (cf. remarque 7.1)

$$i) \quad (B_i u - B_i v, u-v) \geq \alpha \|u - v\|_i^2, \quad \alpha > 0$$

ii) $\forall N_i$ constante positive, $\exists C_i(N_i) > 0$ telle que si $\|u\|_i \leq N_i, \|v\|_i \leq N_i$, alors

$$(B_i u - B_i v, w) \leq C_i(N_i) \|u-v\|_i \|w\|_i \quad \forall w \in V_i.$$

Alors pour i, n, ε_0, k fixés posons $w_i = u^{n+\frac{1}{q}}$.

La suite (w_i^m) définie par l'itération :

$$(7.12) \quad w_i^{m+1} = P_{X_i} (w_i^m - \rho (A_i w_i^m + \frac{\varepsilon_0}{k} w_i^m - g_i))$$

converge fortement vers w_i solution de $(7.9)_i$ pour w_i^0 convenable et

$$\rho = \alpha / C_i^2(N_i) \quad \text{où} \quad g_i = f_i + \frac{\varepsilon_0}{k} u^{n+\frac{i-1}{q}};$$

$$N_i = \frac{2}{k} \|g_i - A_i(0)\|_i \quad (\text{si } w_i^0 = 0)$$

Démonstration

1°) Pour cette assertion, cf. H. Brézis [5].

2°) Comme $\frac{1}{N} \sum_{n=0}^{N-1} u^{n+\frac{1}{q}} \rightarrow u$ dans V fort ; il s'agit maintenant de calculer $u^{n+\frac{1}{q}}$ solution de $(7.9)_i$ pour ε_0, k, n et i fixés.

Posons $B_i = A_i + \frac{\varepsilon_0}{k} I$, $g_i = f_i + \frac{\varepsilon_0}{k} u^{n+\frac{i-1}{q}}$ et $w_i = u^{n+\frac{1}{q}}$

Le problème $(7.9)_i$ est alors équivalent à

$$(7.13)_i \quad (B_i w_i, v-w_i) \geq (g_i, v-w_i) \quad \forall v \in X_i$$

On applique alors à $(7.13)_i$ le théorème 1.1 de [22].

Remarque 7.1

Dans le cas où V et V_i $i = 1, \dots, q$ sont des Banach réflexifs, on peut utiliser des itérations analogues à $(7.12)_i$ moyennant une discrétisation du problème $(7.9)_i$ et ceci de la manière suivante :

1) On se donne un paramètre $h > 0$ destiné à tendre vers 0 et un espace de dimension finie V_{ih} . Comme dans le paragraphe I-5, on remplace le problème $(7.9)_i$ par le problème

$$(7.14)_i \quad \left(\left(A_{ih} + \frac{\varepsilon_0}{k} \right) u_h^{n+\frac{i-1}{q}}, v_h - u_h^{n+\frac{i-1}{q}} \right)_h \geq \left(f_{ih} + \frac{\varepsilon_0}{k} u_h^{n+\frac{i-1}{q}}, v - u_h^{n+\frac{i-1}{q}} \right)_h \quad \forall v_h \in X_h$$

On munit V_{ih} de deux normes équivalentes $\| \cdot \|_{ih}$ et $| \cdot |_{ih}$. Soit $P_{ih} \in \mathcal{L}(V_{ih}, V_i)$ un prolongement de V_{ih} dans V_i injectif. On applique le théorème 5.1 pour montrer que sous les hypothèses du théorème 5.1 on a $P_h u_h^{n+\frac{1}{q}} \longrightarrow u^{n+\frac{1}{q}}$ dans V_i fort quand $h \rightarrow 0$.

2) Il reste maintenant à résoudre par itérations le problème $(7.14)_i$ pour i, n, ϵ_0, k et h fixés. On remplace alors les hypothèses i) et ii) du théorème 7.1 par les suivantes plus générales :

$$\text{iii)} \quad (A_i u - A_i v, u-v) \geq (\varphi(\|u\|_i) - \varphi(\|v\|_i)) (\|u\|_i - \|v\|_i)$$

iv) $\forall N_i$ constante positive $\exists C_i(N) > 0$ telle que

si $\|u\|_i \leq N_i, \|v\|_i \leq N_i$ on ait :

$$(A_i u - A_i v, w) \leq C_i(N_i) \|u-v\|_i \|w\|_i \quad \forall w \in V_i$$

Posons alors $W_{ih} = u_h^{n+\frac{1}{q}}$. La suite (W_{ih}^m) définie par les itérations :

$$(7.15)_i \quad W_{ih}^{m+1} = P_{X_{ih}} (W_{ih}^m - P_{ih} (A_{ih} W_{ih}^m + \frac{\epsilon_0}{k} W_{ih}^m - g_{ih}))$$

$$\text{où } g_{ih} = f_{ih} + \frac{\epsilon_0}{k} u_h^{n+\frac{1}{q}}$$

converge dans V_{ih} (h fixé et $m \rightarrow \infty$) vers $W_{ih} = u_h^{n+\frac{1}{q}}$

solution de $(7.14)_i$ pour W_{ih}^0 convenable et

$$P_{ih} = \frac{\epsilon_0 C_1^2(h)}{k (C_1(N) + \frac{\epsilon_0}{k} C_2^2(h))^2}$$

où $C_1(h)$ et $C_2(h)$ sont telles que

$$C_1(h) \|u_h\|_{ih} \leq |u_h|_{ih} \leq C_2(h) \|u_h\|_{ih}$$

3) Dans les itérations (7.12) et (7.15)_i on peut avoir intérêt à utiliser des paramètres ρ et ρ_i calculés à chaque itération ($\rho = \rho_n$, $\rho_i = \rho_{in}$) cf. [22]

Remarque 7.2 Soit F une fonctionnelle sur V convexe et dérivable au sens de Gateaux. Alors minimiser F sur un convexe fermé X de V de la forme

$X = \bigcap_{i=1}^q X_i$ revient à résoudre une inéquation de la forme (7.1) et le théorème 7.1 s'applique. En effet, le problème

$$(7.16) \quad F(u) \leq F(v) \quad \forall v \in X = \bigcap_{i=1}^q X_i$$

est équivalent au problème :

$$(7.17) \quad (F'(u), v-u) \geq 0 \quad \forall v \in X = \bigcap_{i=1}^q X_i$$

Remarque 7.3

Si le convexe $X = V$, on peut toujours résoudre l'équation

$$(7.18) \quad Au = f \quad \text{où } f \text{ est donné dans } V' \text{ et } A : V \rightarrow V'$$

par décomposition de l'opérateur A et du second membre f . On pose :

$$A = \sum_{i=1}^q A_i \quad \text{et} \quad f = \sum_{i=1}^q f_i$$

on part de $u^0 = 0$; ayant déterminé $u^{n+\frac{1-1}{q}} \in V$ on obtient

$u^{n+\frac{1}{q}} \in V$ comme solution de

$$(7.19) \quad (A_1 + \frac{\epsilon_0}{k}) u^{n+\frac{1}{q}} = f_1 + \frac{\epsilon_0}{k} u^{n+\frac{1-1}{q}}$$

sous les hypothèses (7.8) et les hypothèses sur ϵ_0 et k on a :

$$\frac{1}{N} \sum_{n=0}^{N-1} u^{n+\frac{1}{q}} \rightarrow u \text{ dans } V \text{ fort ;}$$

u étant la solution du problème (8.18). La deuxième partie du théorème 7.1. et la remarque 7.1. s'appliquent encore.

(cf. résultats numériques dans [22]).

La méthode (7.12) suppose que l'on sache projeter par un procédé numérique simple sur le convexe X_i ou que l'on ramène la projection sur X_i à la projection sur X_{i_0} convexe plus simple (cf. [22] théorèmes 2.1; 22). Dans le cas contraire, on aura à résoudre pour $i = i_0$ l'inéquation

$$(7.20) \quad \left(A_{i_0} u^{n+\frac{1_0}{q}} + \frac{\epsilon_0}{k} u^{n+\frac{1_0}{q}}, v-u^{n+\frac{1_0}{q}} \right) \gg \left(f_{i_0} + \frac{\epsilon_0}{k} u^{n+\frac{1_0-1}{q}}, v-u^{n+\frac{1_0}{q}} \right) \quad \forall v \in X_{i_0}$$

pour n , ϵ_0 et k fixés et où X_{1_0} est un convexe "compliqué".
Après un changement de notations évident, résoudre (7.20) équivaut à résoudre une inéquation de la forme

$$(7.21) \quad (B u, v-u) \geq (g, v-u) \quad \forall v \in X$$

où X est un convexe sur lequel on ne sait pas projeter. On fait les hypothèses suivantes :

i) Le convexe X est de la forme

$$X = \{v \in V \mid C v = 0\} \text{ où } C \text{ est un opérateur monotone hémicontinu.}$$

ii) B est un opérateur hémicontinu de V dans V' tel que

$$(B u - B v, u-v) \geq (\varphi(\|u\|) - \varphi(\|v\|)) (\|u\| - \|v\|) \quad \forall u, v \in V$$

avec φ strictement croissante de R_+ dans R tel que $\lim_{r \rightarrow \infty} \varphi(r) = +\infty$.

Soit $\epsilon > 0$ un paramètre destiné à tendre vers 0.

Le théorème 7.1 dit alors que la solution $u_\epsilon \in V$ de l'équation

$$(7.22) \quad B u_\epsilon + \frac{1}{\epsilon} C u_\epsilon = g$$

converge fortement dans V vers u solution de 7.21 quand $\epsilon \rightarrow 0$.

Remarque 7.4

1) Si le convexe X est de la forme

$X = \{v \in V \mid G(v) = 0\}$ avec $G > 0$ convexe s.c.i. et dérivable au sens de Gateaux alors la suite u_ϵ solution de

$$(7.23) \quad B u_\epsilon + \frac{1}{\epsilon} G' u_\epsilon = g$$

converge dans V fort vers u solution de (7.21)

2) Si $X = \{v \in V \mid G_0(v) \leq 0\}$ avec $G_0 : V \rightarrow R$ convexe s.c.i. On pose alors $G = H G_0^2$

$$\text{où } H = \begin{cases} 0 & \text{si } G_0 > 0 \\ 1 & \text{si } G_0 \leq 0 \end{cases}$$

Si G est dérivable au sens de Gateaux alors la suite u_ϵ solution de

$$(7.24) \quad B u_\epsilon + \frac{1}{\epsilon} G' u_\epsilon = g$$

converge dans V fort vers u solution de (7.21).

3) Si $X = \{v \in V \mid G_i(v) \leq 0 \quad i = 1, \dots, p\}$

avec G_i convexe s.c.i ; on introduit alors la fonctionnelle

$$G = \sum_{i=1}^n H_i G_i^2 \quad \text{où} \quad H_i = \begin{cases} 0 & \text{si } G_i > 0 \\ 1 & \text{si } G_i \leq 0 \end{cases}$$

et l'on pose $\varepsilon = (\varepsilon_1, \dots, \varepsilon_p)$ $\varepsilon_i > 0$. Alors la solution u_ε solution de (7.24) converge vers u solution de (7.21) dans V fort quand $\varepsilon \rightarrow 0$.

I-8 Optimisation de fonctionnelle sur un convexe fermé: méthodes directes de résolution.

Soit V un Banach réflexif de dual V' . Soit X un convexe fermé de V et F une fonctionnelle définie sur X à valeurs dans $]-\infty, +\infty]$ non identique à $+\infty$. On suppose que X contient l'origine. On se propose d'approcher la solution $u \in X$ du problème

$$(8.1) \quad F(u) \leq F(v) \quad \forall v \in X$$

Soit $h > 0$ un paramètre destiné à tendre vers 0. On lui associe un espace de dimension V_h de dual V'_h . On munit V_h du produit scalaire $(\cdot, \cdot)_h$ et la norme $\|\cdot\|_h$. Soit $p_h \in \mathcal{L}(V_h, V)$ un prolongement injectif de V_h dans V . On pose $\|u_h\|_h = \|p_h u_h\| \quad \forall u_h \in V_h$. On se donne un convexe X_h fermé de V_h . On cherche maintenant $u_h \in X_h$ tel que

$$(8.2) \quad F(p_h u_h) \leq F(p_h v_h) \quad \forall v_h \in X_h.$$

Nous allons montrer que sous certaines conditions on a $p_h u_p \rightarrow u$ solution de (8.1) quand $h \rightarrow 0$. Plus précisément, on fait les hypothèses suivantes :

i) F est une fonctionnelle strictement convexe telle que

$\lim_{\|u\| \rightarrow +\infty} F(u) = +\infty$; dérivable au sens de Gateaux avec F' borné et vérifiant :

$$(F'u - F'v, u - v) \geq (\varphi(\|u\|) - \varphi(\|v\|)) (\|u\| - \|v\|) \quad \forall u, v \in X$$

où $\varphi: \mathbb{R}_+ \rightarrow \mathbb{R}$ strictement croissante et telle que $\lim_{r \rightarrow \infty} \varphi(r) = +\infty$.

ii) V est un Banach uniformément convexe et $\forall x \in X$, $\exists u_h \in X_h$ tel que $p_h u_h \rightarrow x$ dans V fort.

- iii) F est strictement convexe continu et tel que $\lim_{\substack{\|u\| \rightarrow \infty \\ u \in X}} F(u) = +\infty$
- iv) Les conditions $p_h u_h \rightarrow u$ dans V faible et $F(p_h u_h) \rightarrow F(u)$ entraînent $p_h u_h \rightarrow u$ dans V fort.

Nous avons le

Théorème 8.1

- 1°) On suppose i) et ii) ; alors $\exists u \in X$ unique solution de (8.1). Il existe $u_h \in X_h$ unique solution de (8.2) et $p_h u_h \rightarrow u$ dans V fort.
- 2°) Si on suppose iii) et iv) on a les mêmes conclusions que dans 1°).

Démonstration.

1°) L'existence et l'unicité de $u \in X$ solution de (8.1) et de $u_h \in X_h$ solution de (8.2) résultent du théorème 2.2 car i) $\Rightarrow F$ est continu.

Montrons maintenant que $p_h u_h \rightarrow u$ dans V fort. On montre facilement que $\|p_h u_h\|$ est borné. Donc suivant un ultrafiltre $\mathcal{U}$ on a $p_h u_h \rightarrow \xi$ dans V faible. Comme F' est borné alors $F' p_h u_h \rightarrow \eta$ dans V' faible. On sait que toute solution de (8.1) est solution de :

$$(8.3) \quad (F' p_h u_h, p_h v_h - p_h u_h) \geq 0 \quad \forall v_h \in X_h$$

ce qui donne :

$$(8.4) \quad (F' p_h u_h, p_h u_h) \leq (F' p_h u_h, p_h v_h) \quad \forall v_h \in X_h$$

Puisque $\xi \in X$ d'après ii) $\exists \xi_h \in X_h$ tel que $p_h \xi_h \rightarrow \xi$ dans V fort.

Posons alors $v_h = \xi_h$ dans (8.4) il vient :

$$(8.5) \quad \lim_{\mathcal{U}} \sup (F' p_h u_h, p_h u_h) \leq (\eta, \xi)$$

Donc $F' \xi = \eta$. En passant à la limite dans (8.3) on a d'après ii)

$$(8.6) \quad (F' \xi, v - \xi) \geq 0 \quad \forall v \in X$$

ce qui équivaut à

$$(8.7) \quad F(\xi) \leq F(v) \quad \forall v \in X.$$

D'après l'unicité de la solution on a $\xi = u$.

Donc $p_h u_h \rightarrow u$ dans V faible.

Par ailleurs d'après (8.5) on a :

$$\lim_{\mathcal{U}} \sup (F'(p_h u_h) - F'u, p_h u_h - u) \leq 0$$

$$\text{Donc } \lim_{h \rightarrow 0} (\varphi(\|p_h u_h\|) - \varphi(\|u\|)) (\|p_h u_h\| - \|u\|) = 0$$

Donc $\|p_h u_h\| \rightarrow \|u\|$. Comme V est uniformément convexe et $p_h u_h \rightarrow u$ dans V faible alors $p_h u_h \rightarrow u$ dans V fort suivant $\mathcal{U}$ et suivant tout ultrafiltre plus fin.

2°) Soit v_h^0 fixé dans X_h . Posons

$$K_{v_h^0} = \{u_h \in X_h \mid F(p_h u_h) \leq F(p_h v_h^0)\}$$

On montre facilement que toute solution de (8.2) est solution de

$$(8.8) \quad F(p_h u_h) \leq F(p_h v_h) \quad \forall v_h \in K_{v_h^0}$$

et réciproquement. Par ailleurs, $K_{v_h^0}$ est borné sinon $\exists u_{h_n}$ tel que $\|p_h u_{h_n}\| \rightarrow \infty$ et on aurait $F(p_h u_{h_n}) \rightarrow \infty$ ce qui contredit (8.8).

Donc $\|p_h u_h\|$ est borné et suivant un ultrafiltre $\mathcal{U}$ on a $p_h u_h \rightarrow \xi$ dans V faible. Par ailleurs

$$(8.9) \quad F(\xi) \leq \lim_{\mathcal{U}} \inf F(p_h u_h) \leq F(p_h v_h) \quad \forall v_h \in X_h$$

or $\forall v \in X \exists v_h \in X_h$ tel que $p_h v_h \rightarrow v$

En passant à la limite dans (8.9) on a :

$$(8.10) \quad F(\xi) \leq F(v) \quad \forall v \in X$$

D'après l'unicité de la solution de (8.1) on a $\xi = u$.

Donc $p_h u_h \rightarrow u$ dans V faible. Comme F est continu, d'après iv) on a $p_h u_h \rightarrow u$ dans V fort. D'où le théorème.

Nous allons donner maintenant une méthode numérique permettant de résoudre le problème (8.2), qui peut encore s'énoncer ainsi ; (cf aussi [2'], [11'], [24']) on cherche $x \in X$, convexe fermé de $\mathbb{R}^n$ tel que

$$(8.11) \quad f(x) \leq f(y) \quad \forall y \in X$$

où f est une fonction de n variables à valeurs dans $\mathbb{R}$ de classe $\mathcal{C}^1$ strictement convexe et telle que $\lim_{\substack{\|x\| \rightarrow \infty \\ x \in X}} f(x) = +\infty$.

Approximation de x et $f(x)$: Méthode de la Mire

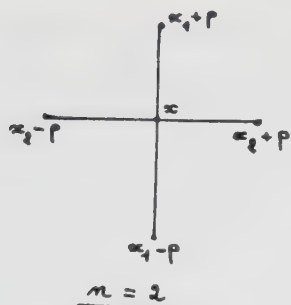
La méthode suivante permet d'approximer x et $f(x)$ vérifiant (8.11)
On suppose dans une première étape que X est un convexe compact de $\mathbb{R}^n$.
On se passera ensuite de cette hypothèse.

Notations

Soit ρ un paramètre positif destiné à tendre vers 0 .
Pour $x \in X$ $x = (x_1, \dots, x_n)$ on note :
 $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, \dots, 0)$; $e_i = (0, \dots, 1, 0, \dots, 0)$

$$M_{x, \rho} = \left\{ \bigcup_{\substack{i=1 \\ \epsilon=0,1,-1}}^n x + \epsilon \rho e_i \right\}$$

C'est la mire de centre x et de rayon ρ . (cf. figure)



On désigne par

$$E(M_{x, \rho}) = \left\{ \bigcup_{\epsilon=-1,1}^n x + \epsilon \rho e_i \right\}$$

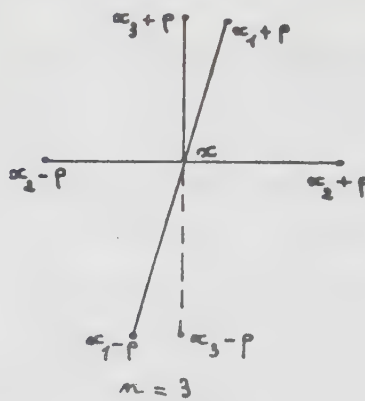
l'ensemble des extrémités de la mire.

Mire de type I

On désigne par $x^p = (x_1^p, \dots, x_n^p)$ le point générique $x = (x_1, \dots, x_n)$ à l'itération p et par ρ_p la valeur du paramètre ρ à l'itération p .

On dira que M_{x^p, ρ_p} est de type I, s'il existe $y \in E(M_{x^p, \rho_p})$ tel que

$f(y) < f(x^p)$ (c'est-à-dire si le minimum sur la mire est un point de l'extrémité).



Mire de type II

On dira que $M_{x^p, p}$ est de type II si $\forall y \in L(M_{x^p, p})$

On a : $f(y) \geq f(x^p)$

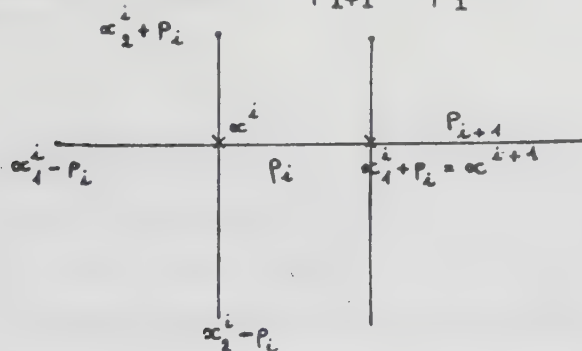
(c'est-à-dire le minimum est au centre de la mire)

Soient $x^0 \in X$ et $p_0 > 0$. On leur associe une première mire M_{x^0, p_0} .

On définit par récurrence la suite M_{x^n, p_n} de la manière suivante :

1) Si M_{x^i, p_i} est de type I, on pose x^{i+1} = l'un des points de $E(M_{x^i, p_i})$

où le minimum est atteint et $p_{i+1} = p_i$.



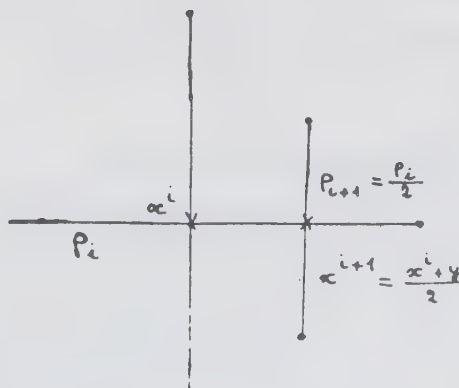
2) Si M_{x^i, p_i} est de type II alors soit Γ_i la mire dont l'une des extrémités est x^i ; une autre étant l'un des points $y \in E(M_{x^i, p_i})$ où f atteint son minimum sur $L(M_{x^i, p_i})$. Le rayon de Γ_i est alors $p = \frac{p_i}{2}$

Nous avons alors deux cas

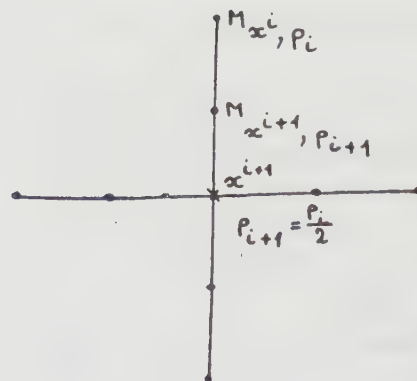
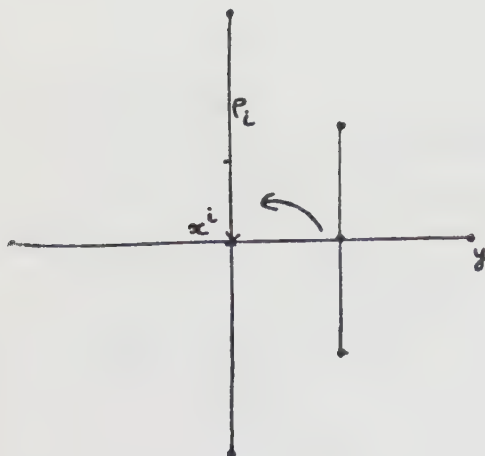
α) Si Γ_i est de type I on pose

$$x^{i+1} = \frac{x^i + y}{2} \quad \text{et} \quad p_{i+1} = \frac{p_i}{2}$$

ce qui définit $M_{x^{i+1}, p_{i+1}}$.



$\beta)$ Si Γ_1 est de type II, on pose alors $M_{x^{i+1}, p_{i+1}} = M_{x^i, \frac{p_i}{2}}$



La technique consiste donc en ceci : si M_{x^i, p_i} est de type I on poursuit le procédé et au bout d'un nombre fini d'itérations (car le convexe X est borné) la mire devient de type II qu'on quitte en deux coups après avoir diminué le paramètre p . On reprend alors la recherche du minimum avec une mire de type I.

Si certains points de la mire M_{x^i, p_i} n'appartiennent pas au convexe X , on n'en tiendra pas compte même si f est défini sur tout $\mathbb{R}^n$.

On désigne maintenant par C_n la suite infinie des mires de type II de centre x^n et de rayon p_n .

Nous avons alors le

Théorème 8.2

Soit X un convexe fermé de $\mathbb{R}^n$. Soit $f : X \rightarrow \mathbb{R}$ strictement convexe, s.c.i. et tel que $\lim_{\substack{\|x\| \rightarrow \infty \\ x \in X}} f(x) = +\infty$.

Soit $x \in X$ unique solution de

$$(8.12) \quad f(x) \leq f(y) \quad \forall y \in X.$$

On désigne par C_p la suite des mires de type II définie plus haut ($\rho_p \rightarrow 0$ quand $p \rightarrow \infty$). Si f est définie sur un voisinage compact de X et de classe C^1 , alors la suite (ξ^p) converge vers x solution de (8.12) quand $p \rightarrow \infty$. De plus $f(x) = \inf_{y \in X} f(y)$

Démonstration

Voici une démonstration dans le cas où $X = \mathbb{R}^n$
a) Soit $y_0 \in \mathbb{R}^n$ fixé. Posons :

$$K_{y_0} = \{ y \in X \mid f(y) \leq f(y_0) \}$$

alors K_{y_0} est un convexe fermé et borné de $\mathbb{R}^n$ et toute solution de (8.12) est solution de

$$(8.13) \quad f(x) \leq f(y) \quad \forall y \in K_{y_0}$$

et réciproquement.

On est donc ramené au cas où la solution se trouve à l'intérieur du convexe fermé K_{y_0} de $\mathbb{R}^n$. (Si le minimum est sur la frontière alors il est confondu avec y_0)

b) La suite (ξ^p) étant bornée, on peut extraire une sous suite encore notée $(\xi^p) \rightarrow \xi = (\xi_1, \dots, \xi_n)$. Par ailleurs, on a pour ρ_p suffisamment petit :

$$f(\xi_1^p, \dots, \xi_1^p + \rho_p, \dots, \xi_n^p) = f(\xi_1^p, \dots, \xi_1^p, \dots, \xi_n^p) + \rho_p f'_{\xi_1^p}(\xi_1^p, \dots, \xi_1^{p+\theta_1^p} \rho_p, \dots, \xi_n^p)$$

avec $0 < \theta_1^p(\rho_p) < 1$

$$f(\xi_1^p, \dots, \xi_1^p - \rho_p, \dots, \xi_n^p) = f(\xi_1^p, \dots, \xi_1^p, \dots, \xi_n^p) + \rho_p f'_{\xi_1^p}(\xi_1^p, \dots, \xi_1^{p-\theta_1^p} \rho_p, \dots, \xi_n^p)$$

avec $0 < \theta_1^p(\rho_p) < 1$

ce qui donne

$$f'(\xi_1^p + \theta_1^p \rho_p) \geq 0 \quad \text{et} \quad f'(\xi_1^p - \theta_1^p \rho_p) \leq 0 \quad \text{pour } i = 1, \dots, n.$$

Donc $f'(\xi) = 0$. Comme le minimum est unique, il résulte que $\xi = x$

et la suite (ξ^p) entière converge vers x vérifiant $f(x) \leq f(y) \quad \forall y \in X$.

Remarque pratique

Si le convexe X est trop effilé, il y a lieu de faire une rotation des axes avant de démarrer la méthode. Dans certains cas, ceci accélère nettement la convergence. Au stade de l'exploitation numérique de cette méthode nous avons une "stratégie" d'exécution qui donne le résultat en un temps très nettement inférieur à celui de la méthode brute.

Chapitre II

APPLICATIONS AUX EQUATIONS ET INEQUATIONS AUX DERIVEES PARTIELLES

II-1 Résolution d'une inéquation variationnelle non linéaire p. 42

1.1 Existence et unicité de la solution p. 42

1.2 Méthode d'approximation par régularisation p. 43

1.3 Problème de minimisation correspondant; autres exemples p. 44

1.4 Méthode d'approximation par discrétisation p. 47

II-2 Une généralisation non linéaire de l'opérateur du second ordre p. 48

II-3 Etude d'une équation non linéaire p. 58

II-4 Etude d'un problème de mécanique p. 60

II-5 Applications de la méthode de la miré p. 72

Problème 1 p. 72

Problème 2 p. 73

Problème 3 p. 74

II-1. RESOLUTION D'UNE INEQUATION VARIATIONNELLE NON LINEAIRE.

Soit Ω un ouvert borné de R^n de frontière $\Gamma = \partial \Omega$ suffisamment régulière. On note par $x = (x_1, \dots, x_n)$ un point de R^n . On désigne par $W^{1,p}(\Omega)$, $1 < p < +\infty$ l'espace des (classes de) fonctions $u \in L^p(\Omega)$ telles que $D_i u = \frac{\partial u}{\partial x_i} \in L^p(\Omega)$; les dérivées étant prises au sens des distributions. $W_0^{1,p}(\Omega)$ désigne l'adhérence de $\mathcal{D}(\Omega)$ dans $W^{1,p}(\Omega)$.

On note par $W^{-1,q}(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$, le dual de $W_0^{1,p}(\Omega)$.

Posons $V = W_0^{1,p}(\Omega)$, On sait (cf. [15]) que, V coïncide avec l'espace $\{u \mid u \in L^p(\Omega), D_i u \in L^p(\Omega), i = 1, \dots, n; u|_{\Gamma} = 0\}$

et $V' = W^{-1,q}(\Omega) = \{f \mid f = f_0 + \sum_{i=1}^n D_i f_i, f_0, f_i \in L^q(\Omega)\}$.

Muni de la norme $\|u\| = \left(\sum_{i=1}^n \|D_i u\|_{L^p(\Omega)}^p \right)^{1/p}$, V est un espace uniformément convexe.

Soit X un convexe de V contenant l'origine.

On pose : $A u = - \sum_{i=1}^n D_i (|D_i u|^{p-2} D_i u)$ $p \geq 2$

et l'on cherche $u \in X$ tel que

$$(1.1) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

pour f donné dans V' .

1.1 - Existence et unicité de la solution.

Nous avons le

Théorème 1.1

Il existe $u \in X$ unique solution de (1.1)

Démonstration

L'application $t \longrightarrow (A(tu + (1-t)v), w)$ est continue

$\forall u, v, w \in V$. De plus on a :

$$(Au, u) = \|u\|_{L^p(\Omega)}^p$$

$$|(Au, v)| = \left| \sum_{i=1}^n (|D_i u|^{p-2} D_i u, D_i v) \right| \leq \sum_{i=1}^n \|D_i u\|_{L^p(\Omega)}^{p-1} \|D_i v\|_{L^p(\Omega)}$$

Donc

$$(Au, v) \leq \|u\|^{p-1} \|v\|$$

ce qui donne :

$$(Au - Av, u - v) \geq (\|u\|^{p-1} - \|v\|^{p-1}) (\|u\| - \|v\|)$$

D'après le théorème 1.1 chap. I on a l'existence et l'unicité de $u \in X$ solution de (1.1).

1.2 - Méthode d'approximation par régularisation

On démontre de même le théorème suivant :

Théorème 1.2

Soit X un convexe fermé de $W_0^{1,p}(\Omega)$ contenant l'origine

Soit B un opérateur de $W_0^{1,p}(\Omega)$ dans $W^{-1,q}(\Omega)$ monotone hémicontinu. On pose :

$$A_\lambda u = - \sum_{i=1}^n D_i (|D_i u|^{p-2} D_i u) + \lambda B u \quad \lambda \geq 0 \quad \text{Alors}$$

1 - Pour tout $f \in W_0^{1,q}(\Omega)$, il existe $u_\lambda \in X$ unique solution de

$$(1.2) \quad (A_\lambda u_\lambda, v - u_\lambda) \geq (f, v - u_\lambda) \quad \forall v \in X.$$

2 - Si B vérifie de plus

$$(Bu - Bv, u - v) \geq (\varphi(\|u\|) - \varphi(\|v\|))(\|u\| - \|v\|)$$

avec φ strictement croissante de R_+ dans R , telle que

$\lim_{r \rightarrow \infty} \varphi(r) = +\infty$; alors la solution u_λ de (1.2) converge fortement dans $W_0^{1,p}(\Omega)$ quand $\lambda \rightarrow 0$ vers u solution de (1.1).

1.3 - Problème de minimisation correspondant ; autres exemples

Nous allons maintenant préciser les problèmes de minimisation sur le convexe X correspondants aux inégalités variationnelles (1.1) et (1.2).

Théorème 1.3

Soit $F(u) = \frac{1}{p} \sum_{i=1}^n \|D_i u\|_{L^p(\Omega)}^p$, $p \geq 2$. Alors F est

une fonctionnelle convexe sur $W_0^{1,p}(\Omega)$ dérivable au sens de Gateaux et

$$F'(u) = - \sum_{i=1}^n D_i (|D_i u|^{p-2} D_i u)$$

Nous utiliserons le lemme suivant :

Lemme 1.1

Soit $G(u) = \frac{1}{p} \|u\|_{L^p(\Omega)}^p$, $p \geq 2$ et q donné par

$\frac{1}{p} + \frac{1}{q} = 1$. Alors G est une fonction convexe sur $L^p(\Omega)$

dérivable au sens de Gateaux et $G' : L^p(\Omega) \rightarrow L^q(\Omega)$ est définie

par $G'(u) = |u|^{p-2} u$.

Démonstration

Il suffit de montrer que

$$(G'(u), v) = \int_{\Omega} |u|^{p-2} u v \, dx = \frac{1}{p} \lim_{t \rightarrow 0} \frac{\|u + tv\|^p - \|u\|^p}{t}$$

$$\text{Mais } \frac{\|u + tv\|^p - \|u\|^p}{t} = \int_{\Omega} \frac{|u + tv|^p - |u|^p}{t} \, dx$$

On montre facilement que l'application

$$t \longrightarrow h(t) = \frac{|a + tb|^p - |a|^p}{t} \quad t \in]0, 1] ; \quad a, b \in \mathbb{R} ; \quad p > 2,$$

est croissante de t et $\lim_{t \rightarrow 0} h(t) = p|a|^{p-2} ab$

Ce qui donne

$$\frac{|u(x) + tv(x)|^p - |u(x)|^p}{t} \leq |u(x) + v(x)|^p - |u(x)|^p \quad \forall t \in]0, 1]$$

D'après le théorème de la convergence dominée de Lebesgue, on a :

$$\lim_{t \rightarrow 0} \frac{\|u + tv\|^p - \|u\|^p}{t} = \int_{\Omega} |u|^{p-2} u v \, dx. \quad \text{D'où le lemme.}$$

Démonstration du théorème

Soient $u, v \in W_0^{1,p}(\Omega)$ alors $D_i u \in L^p(\Omega)$, $D_i v \in L^p(\Omega)$

pour $i = 1, \dots, n$. Il suffit alors de montrer que

$$\begin{aligned} (F'(u), v) &= \sum_{i=1}^n \int_{\Omega} |D_i u|^{p-2} D_i u D_i v \, dx \\ &= \frac{1}{p} \lim_{t \rightarrow 0} \sum_{i=1}^n \frac{\|D_i u + t D_i v\|_{L^p(\Omega)}^p - \|D_i u\|_{L^p(\Omega)}^p}{t} \end{aligned}$$

Ce que l'on a d'après le lemme 1.1

Soit A une application de $W_0^{1,p}(\Omega)$ dans $W^{-1,q}(\Omega)$ définie par

$$(1.3) \quad A u = - \sum_{i=1}^n D_i (|D_i u|^{p-2} D_i u) \quad p \geq 2$$

Soient X un convexe fermé de $W_0^{1,p}(\Omega)$ et f donné dans $W^{-1,q}(\Omega)$

Posons alors

$$(1.4) \quad J(u) = \frac{1}{p} \sum_{i=1}^n \|D_i u\|_{L^p}^p - (f, u)$$

$$(1.5) \quad A_1 u = A u - \lambda \Delta u + \mu u \quad \lambda, \mu \geq 0 ; \quad p \geq 2$$

$$(1.6) \quad J_1(u) = J(u) + \frac{\lambda}{2} \sum_{i=1}^n \|D_i u\|_{L^2}^2 + \frac{\mu}{2} \|u\|_{L^2}^2 - (f, u)$$

Nous avons alors le

Corollaire 1.1

1 - Toute solution $u \in X$ de

$$(1.7) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

est solution de

$$(1.8) \quad J(u) \leq J(v) \quad \forall v \in X$$

et réciproquement

De plus on a l'existence et l'unicité de la solution $u \in X$

2 - On a la même assertion en remplaçant A par A_1 et J par J_1 .

1.4 Méthode d'approximation par discrétisation.

Nous allons maintenant approcher la solution $u \in X$ de l'inéquation

$$(1.9) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X; f \text{ donné dans } W^{-1,p}(\Omega)$$

avec

$$Au = - \sum_{i=1}^n D_i (|D_i u|^{p-2} D_i u) - \lambda \Delta u.$$

Posons

$$W^{1,p}_i(\Omega) = \{ \varphi \mid \varphi \in L^p(\Omega), D_i \varphi \in L^p(\Omega) \}$$

Dans [7] nous donnons d'après [1] un espace de dimension fini

$V_h(\Omega_h)$ qui approche $W^{1,p}_0(\Omega)$. Nous donnons de plus un prolongement injectif $p_h^i \in \mathcal{L}(V_h(\Omega_h), W^{1,p}_i(\Omega))$. Il est tel que

$$D_i p_h^i u_h = \nabla_i u_h \quad \text{pour } u_h \in V_h(\Omega_h).$$

$$\text{où } \nabla_i u_h(x) = \frac{u_h(x + \frac{h_i}{2}) - u_h(x - \frac{h_i}{2})}{h_i}$$

Ce qui donne :

$$\begin{aligned} (A_h u_h, u_h)_h &= \sum_{i=1}^n \int_{\Omega} |D_i p_h^i u_h|^{p-2} D_i p_h^i u_h D_i p_h^i v_h \, dx \\ &\quad + \lambda \int_{\Omega} \sum_{i=1}^n D_i p_h^i u_h D_i p_h^i v_h \, dx \end{aligned}$$

$$(A_h u_h, v_h)_h = \sum_{i=1}^n \int_{\Omega} |\nabla_i u_h|^{p-2} \nabla_i u_h \nabla_i v_h \, dx + \lambda \sum_{i=1}^n \int_{\Omega} \nabla_i u_h \nabla_i v_h \, dx$$

D'où

$$(1.10) \quad A_h u_h = - \sum_{i=1}^n \nabla_i (|\nabla_i u_h|^{p-2} \nabla_i u_h) - \lambda \sum_{i=1}^n \nabla_i^2 u_h$$

On sait cf [1] qu'il existe un prolongement

$p_h \in \mathcal{L}(V_h(\Omega_h), W_0^{1,p}(\Omega_h))$. Posons alors

$$(1.11) \quad X_h = P_h^{-1}(X)$$

$$f_h = \frac{1}{h_1 \dots h_n} \int_{\Omega} f \theta_h dx \quad \text{où } \theta_h \text{ est la fonction}$$

caractéristique d'un pavé de centre x et de côtés $h_1, \dots, h_n$ respectivement parallèles aux axes de coordonnées $ox_1, \dots, ox_n$.

Le problème (1.9) se discrétise alors sous la forme :

$$(1.12) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h ;$$

et on a le

Théorème 1.4

Le problème (1.12) admet une solution unique $u_h \in X_h$.

De plus $P_h^i u \xrightarrow{h} u$ dans $W_0^{1,p}(\Omega)$ fort pour $i = 0, 1, \dots, n$; u étant la solution de (1.9).

Dans [22] nous étudions la résolution numérique, à l'aide de méthodes itératives, du problème (1.9) et (1.12); pour les convexes suivants:

$$(1.13) \quad X_1 = \{u | u \in W_0^{1,p}(\Omega) ; \sum_{i=1}^n \|D_i u\|_{L^2(\Omega)}^2 \leq 1\}$$

$$(1.14) \quad X_2 = \{u | u \in W_0^{1,p}(\Omega) ; u \geq 0 \text{ p.p sur } E \subset \Omega\}$$

$$(1.15) \quad X_3 = \{u | u \in W_0^{1,p}(\Omega) ; |\text{grad } u| \leq 1 \text{ p.p sur } \Omega\}$$

II-2 UNE GENERALISATION NON LINEAIRE DE L'OPERATEUR DU SECOND ORDRE.

Soit Ω un ouvert borné de R^n . On se donne un opérateur A de $W_0^{1,p}(\Omega)$ dans $W^{-1,q}(\Omega)$ avec $\frac{1}{p} + \frac{1}{q} = 1$ défini par

$$(2.1) \quad A u = - \sum_{i,j=1}^n D_j \left\{ a_{ij} D_i u \left(\sum_{i,j} a_{ij} D_i u D_j u \right)^{\frac{p-2}{2}} \right\} \quad p \geq 2$$

On suppose :

$$(2.2) \quad a_{i,j} \in L^\infty(\Omega) \quad , \quad a_{i,j} = a_{j,i} \quad \text{et} \quad \sum_{i,j} a_{i,j} \xi_i \xi_j \geq \alpha \sum_{i=1}^n |\xi_i|^2$$

$$\alpha > 0 \quad \text{et} \quad \xi_i \in \mathbb{R}$$

Posons

$$(2.3) \quad X = \{ u \in W_0^{1,p}(\Omega) \mid |\text{grad } u| \leq 1 \text{ p.p. sur } \Omega \}$$

On cherche $u \in X$ solution de

$$(2.4) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X \quad \text{pour } f \text{ donné dans } W^{1,q}(\Omega) .$$

Nous avons le

Théorème 2.1

On suppose (2.1), (2.2), (2.3) alors

1 - Toute solution de (2.4) est solution de

$$(2.5) \quad J(u) \leq J(v) \quad \forall v \in X$$

avec

$$J(v) = \frac{2}{p} \int_{\Omega} \left(\sum_{i,j} a_{i,j} D_i v D_j v \right)^{\frac{p}{2}} dx - (f, v)$$

et réciproquement

2 - Le problème (2.4) admet une solution unique.

Auparavant, nous démontrons le

Lemme 2.1

Soient $a, b, c, p \in \mathbb{R} ; \quad p \geq 2$. Alors la fonction

$$t \longrightarrow h(t) = \frac{|a + tb + t^2c|^p - |a|^p}{t} \quad t \in]0, 1] \text{ est croissante}$$

et $\lim_{t \rightarrow 0} h(t) = p|a|^{p-2} ab$.

Démonstration

La fonction $t \longrightarrow |a + t b + t^2 c|^p$ est convexe. Donc $h(t)$ est croissante.

Si $a = 0$ on a $\lim_{t \rightarrow 0} h(t) = 0$

Si $a \neq 0$ on a
$$h(t) = |a|^p \frac{(|1 + t \frac{b}{a} + t^2 \frac{c}{a}|^p - 1)}{t}$$

Quand $t \longrightarrow 0$; $h(t) \sim |a|^p p \left(\frac{b}{a} + t \frac{c}{a} \right)$.

D'où $\lim_{t \rightarrow 0} h(t) = p |a|^{p-2} a b$.

Démonstration du théorème 2.1

1 - Montrons que l'opérateur F défini par

$$F(u) = \frac{2}{p} \int_{\Omega} \left(\sum_{i,j=1}^n a_{ij} D_i u D_j u \right)^{\frac{p}{2}} dx$$

vérifie les assertions suivantes :

α) F applique $W_0^{1,p}(\Omega)$ dans R .

β) F est convexe

γ) F est dérivable au sens de Gateaux et $F' = A$.

α) Soit $a_{ij} \in L(\Omega)$. Il suffira de montrer que $\int_{\Omega} (D_i u D_j u)^{\frac{p}{2}} dx < +\infty$.

Or pour $u \in W_0^{1,p}(\Omega)$ on a :

$$\int_{\Omega} (D_i u D_j u)^{\frac{p}{2}} dx \leq \left(\int_{\Omega} |D_i u|^p dx \right)^{1/2} \left(\int_{\Omega} |D_j u|^p dx \right)^{1/2} < +\infty$$

β) Les applications suivantes sont convexes

$$\xi \in R^n \longrightarrow \sum_{i,j=1}^n a_{ij} \xi_i \xi_j \in R ;$$

$$\sum_{i,j=1}^n a_{i,j} \xi_i \xi_j \longrightarrow \left(\sum_{i,j=1}^n a_{i,j} \xi_i \xi_j \right)^{\frac{p}{2}}$$

Donc $F(u)$ est convexe.

γ) Nous avons successivement :

$$\frac{F(u+tv) - F(u)}{t} = \frac{2}{pt} \int_a^b \left\{ \left[\sum_{i,j=1}^n a_{i,j} D_i(u+tv) D_j(u+tv) \right]^{\frac{p}{2}} - \left[\sum_{i,j=1}^n a_{i,j} D_i u D_j u \right]^{\frac{p}{2}} \right\} dx$$

Posons

$$Y_1 - Y_2 = \left[\sum_{i,j=1}^n a_{i,j} D_i(u+tv) D_j(u+tv) \right]^{\frac{p}{2}} - \left[\sum_{i,j=1}^n a_{i,j} D_i u D_j u \right]^{\frac{p}{2}}$$

on a

$$Y_1 - Y_2 \leq \left[\sum_{i,j=1}^n a_{i,j} D_i u D_j u + 2 \sum_{i,j=1}^n a_{i,j} D_i u D_j v + \sum_{i,j=1}^n a_{i,j} D_i v D_j v \right]^{\frac{p}{2}} - \left[\sum_{i,j=1}^n a_{i,j} D_i u D_j u \right]^{\frac{p}{2}}$$

D'après le théorème de la convergence dominée de Lebesgue on a :

$$\lim_{t \rightarrow 0} \frac{F(u+tv) - F(u)}{t} = \frac{2}{p} \int_a^b \lim_{t \rightarrow 0} \frac{Y_1 - Y_2}{t} dx$$

D'après le lemme 2.1 nous avons :

$$\lim_{t \rightarrow 0} \frac{F(u+tv) - F(u)}{t} = \frac{2}{p} \int_a^b \left| \sum_{i,j=1}^n a_{i,j} D_i u D_j u \right|^{\frac{p-2}{2}} \left(\sum_{i,j=1}^n a_{i,j} D_i u D_j u \right) \left(\sum_{i,j=1}^n a_{i,j} D_i u D_j v \right) dx$$

Donc nous avons au sens des distributions

$$(F'(u), v) = \left\langle \left(\sum_{i,j=1}^n a_{i,j} D_i u D_j u \right)^{\frac{p-2}{2}}, \sum_{i,j=1}^n a_{i,j} D_i u D_j v \right\rangle$$

$$(F'(u), v) = \left\langle - \sum_{i,j=1}^n D_j \left\{ a_{i,j} D_i u \left(\sum_{i,j=1}^n a_{i,j} D_i u D_j u \right)^{\frac{p-2}{2}} \right\}, v \right\rangle$$

D'où $F'(u) = A u$. Le théorème 2.1 chap. I donne alors l'équivalence entre (2.4) et (2.5).

2 - Posons $\nabla u = \left(\sum_{i=1}^n |D_i u|^2 \right)^{1/2}$ on munit alors $W_0^{1,p}(\Omega)$

de la norme $\|u\| = \|\nabla u\|_{L_p}$. Nous avons alors :

$$F(u) = \frac{2}{p} \int_{\Omega} \left(\sum_{i,j} a_{ij} D_i u D_j u \right)^{\frac{p}{2}} dx \geq \frac{2}{p} C^{\frac{p}{2}} \int_{\Omega} \left(\sum_i |D_i u|^2 \right)^{\frac{p}{2}} dx$$

Donc $\frac{F(u)}{\|u\|} \geq C_1 \|u\|^{p-1}$ avec $C_1 > 0$ et $p \geq 2$.

D'après le théorème 2.2 chap. I nous avons alors l'existence et l'unicité de la solution.

Cas particulier

Si $p = 2$; nous avons alors $F(u) = \int_{\Omega} \sum_{i,j=1}^n a_{ij} D_i u D_j u dx$.

Toute solution $u \in X \subset H_0^1(\Omega)$ de

$$(2.6) \quad J(u) \leq J(v) \quad \forall v \in X \quad \text{où} \quad J(v) = F(v) - (f, v)$$

est solution de

$$(2.7) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X \quad \text{pour } f \text{ donné dans } H^{-1}(\Omega)$$

$$\text{où} \quad Au = - \sum_{i,j=1}^n D_j (a_{ij} D_i u) ; \quad A : H_0^1(\Omega) \longrightarrow H^{-1}(\Omega)$$

et réciproquement.

Nous allons donner maintenant une autre méthode pour résoudre (2.7).

On suppose : $a_{ij} \in L^{\infty}(\Omega)$; $a_{ij} = a_{ji}$;

$$\lim_{\|\xi\| \rightarrow \infty} \frac{a_{ij} \xi_i \xi_j}{\|\xi\|} = +\infty, \quad \forall \xi \in \mathbb{R}^n ; \quad \|\xi\| = \left(\sum_{i=1}^n |\xi_i|^2 \right)^{1/2}$$

On munit $H_0^{-1}(\Omega)$ de la norme : $\|u\| = \left(\sum_{i=1}^n \|D_i u\|_{L^2}^2 \right)^{1/2}$

on note par T et Λ les matrices de termes généraux (a_{ij}) et (λ_{ij}) respectivement.

Nous avons le

Théorème 2.2

Soit f donné dans $H^{-1}(\Omega)$ alors

1 - Toute solution de (2.7) est solution de

$$(2.8) \quad J(u) \leq J(v) \quad \forall v \in X$$

$$\text{où} \quad J(v) = \sum_{i=1}^n \frac{1}{2} \left\| \sum_{j=1}^n \lambda_{ij} D_j v \right\|_{L^2}^2 - (f, v)$$

et réciproquement.

Les λ_{ij} sont donnés par l'équation

$$(2.9) \quad T = \Lambda^t \Lambda \quad \Lambda^t \text{ étant la matrice transposée de } \Lambda.$$

2 - On a l'existence et l'unicité de la solution.

On démontre le lemme suivant ;

Lemme 2.2

Si T est une matrice d'ordre n à éléments réels, est définie positive ; alors il existe une matrice Λ (réelle d'ordre n) triangulaire supérieure et non singulière telle que

$$T = \Lambda^t \Lambda$$

Si on choisit $\lambda_{ii} > 0$ la décomposition est unique.

Démonstration du théorème 2.2

$$1 - F(u) = \frac{1}{2} \sum_{i=1}^n \left\| \sum_{j=1}^n \lambda_{ij} D_j u \right\|_L^2 \quad \text{est strictement convexe.}$$

Montrons que F est dérivable au sens de Gateaux et $F' = A$.

Nous avons en effet :

$$\lim_{t \rightarrow 0} \frac{F(u + tv) - F(u)}{t} = \sum_{i=1}^n \left(\sum_{j=1}^n \lambda_{ij} D_j u, \sum_{j=1}^n \lambda_{ij} D_j v \right)$$

Donc

$$(F'(u), v) = \sum_{i=1}^n \left(\sum_{j,k=1}^n \lambda_{ij} \lambda_{ik} D_i u, D_k v \right) \quad \forall v \in X$$

Comme

$$(A u, v) = \sum_{i=1}^n \left(\sum_{k=1}^n a_{ik} D_i u, D_k v \right) \quad \text{et} \quad a_{ik} = a_{ki} ;$$

Nous avons : $F' = A$; si, et seulement si on a :

$$T = A^t A ; \quad \text{avec} \quad T = (a_{ij}) .$$

Comme T est définie positive ; une telle décomposition existe d'après le lemme 2.2. . Le théorème 2.1 chap. I donne alors l'équivalence entre (2.7) et (2.8).

2 - $F(u)$ est strictement convexe, s. c. i

$$\text{De plus} \quad \frac{F(u)}{\|u\|} = \frac{1}{2} \sum_{i=1}^n \frac{\left\| \sum_{j=1}^n \lambda_{ij} D_j u \right\|_L^2}{\|u\|}$$

Posons $\xi_j = D_j u$ il vient :

$$\frac{F(u)}{\|u\|} = \frac{1}{2} \sum_{j=1}^n \frac{\left\| \sum_{j=1}^n \lambda_{1,j} \xi_j \right\|_L^2}{\|\xi\|} = \frac{1}{2} \sum_{j=1}^n \frac{\left(\sum_{j,k=1}^n \lambda_{1,j} \lambda_{1,k} \xi_j, \xi_k \right)}{\|\xi\|}$$

Donc

$$\frac{F(u)}{\|u\|} = \frac{1}{2} \sum_{j=1}^n \frac{(a_j, \xi_j, \xi_j)}{\|\xi\|}$$

Ce qui donne $\lim_{\|u\| \rightarrow \infty} \frac{F(u)}{\|u\|} = +\infty$. D'où l'existence et l'unicité de la solution.

Nous allons nous placer dans le cadre du paragraphe 5 chap. I pour approcher la solution u du problème (2.4) ou (2.5).

Posons :

$$W^{1,p}_t(\Omega) = \{ \varphi \mid \varphi \in L^p(\Omega), D_t \varphi \in L^p(\Omega) \quad t = 1, \dots, n \}$$

Comme dans le paragraphe II-1, 1.4 on se donne un espace de dimension finie V_h ; un prolongement $p_h^t \in \mathcal{L}(V_h, W^{1,p}_t(\Omega))$ injectif et tel que $D_t p_h^t u_h = \nabla_t u_h \quad \forall u_h \in V_h$ pour $t = 1, \dots, n$.

où $\nabla_t u_h(x) = \frac{u_h(x + \frac{h_t}{2}) - u_h(x - \frac{h_t}{2})}{h_t}$

Posons :

$$f_h = \frac{1}{h_1 \dots h_n} \int_{\Omega} f \theta_h^M dx$$

$$a_{1,j}^h(x) = \frac{1}{h_1 \dots h_n} \int_{\Omega} a_{1,j} \theta_h^M dx$$

$$X_h = \{ u_h \in V_h \mid \left(\sum_{i=1}^n |\nabla_i u_h|^2 \right)^{1/2} \leq 1 \}.$$

Nous avons alors le

Théorème 2.3

1 - Il existe $u_h \in X_h$ tel que

$$(2.10) \quad (A_h u_h, v_h - u_h) \geq (f_h, v_h - u_h) \quad \forall v_h \in X_h$$

$$\text{où } A_h u_h = - \sum_{i,j=1}^n \nabla_j \left\{ a_{ij}^h \nabla_i u_h \left(\sum_{i,j=1}^n a_{ij}^h \nabla_i u_h \nabla_j u_h \right)^{\frac{p-2}{2}} \right\} \quad p \geq 2$$

2 - De plus $R_h^1 u_h \longrightarrow u$ dans $W_0^{1,p}(\Omega)$ fort u étant la solution de (2.4) ou de (2.5)

Auparavant, nous allons démontrer le

Lemme 2.3

On se donne $a_{ij} \in L^\infty(\Omega)$ tels que $a_{ij} = a_{ji}$.

Posons le produit scalaire

$$[u, v] = \sum_{i,j=1}^n a_{ij} D_i u D_j v$$

et la norme induite

$$[u] = \left(\sum_{i,j=1}^n a_{ij} D_i u D_j u \right)^{1/2}$$

Soit $\| \cdot \|$ la norme sur $W_0^{1,p}(\Omega)$ définie par

$$\|u\| = \left(\int_{\Omega} [u]^p \right)^{1/p}$$

Alors l'opérateur défini par

$$(2.11) \quad A u = - \sum_{i,j=1}^n D_j \left\{ a_{ij} D_i u \left(\sum_{i,j=1}^n a_{ij} D_i u D_j u \right)^{\frac{p-2}{2}} \right\} \quad p \geq 2$$

vérifie :

$$(2.12) \quad (Au - Av, u - v) \geq (\|u\|^{p-1} - \|v\|^{p-1}) (\|u\| - \|v\|)$$

Démonstration

Nous avons successivement :

$$(Au, u) = \int_{\Omega} \left(\sum_{i,j} a_{i,j} D_i u D_j u \right)^{\frac{p}{2}} dx = \int_{\Omega} [u]^p dx = \|u\|^p$$

$$(Au, v) = \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} D_i u D_j u \right)^{\frac{p-2}{2}} \sum_{i,j=1}^n a_{i,j} D_i u D_j v dx = \int_{\Omega} [u]^{p-2} [u, v] dx$$

$$(Au, v) \leq \int_{\Omega} [u]^{p-1} [v] dx \leq \left(\int_{\Omega} [v]^p dx \right)^{1/p} \left(\int_{\Omega} [u]^{(p-1)q} dx \right)^{1/q} = \|v\| \|u\|^{p-1}$$

D'où

$$(Au - Av, u - v) \geq (\|u\|^{p-1} - \|v\|^{p-1}) (\|u\| - \|v\|)$$

Remarque 2.1

Du lemme précédent et du théorème 1.1 chap. I, on déduit facilement une démonstration directe de l'existence et l'unicité de la solution

$u \in X \subset W_0^{1,p}(\Omega)$ vérifiant :

$$(2.13) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X \text{ pour } f \text{ donné dans } W^{-1,q}(\Omega).$$

Démonstration du théorème 2.3

Nous allons vérifier les hypothèses du théorème 5.1 chap. I. L'opérateur A est monotone hémicontinu et vérifie

$$\lim_{\|u\| \rightarrow \infty} \frac{(Au, u)}{\|u\|} = +\infty ; \quad \text{d'après le lemme précédent. Toujours d'après}$$

le lemme 2.3; si $(Au - Av, u - v) = 0$ alors on a :

$$\|u\| = \|v\|. \text{ De plus, } A \text{ est un opérateur borné.}$$

Puisque $X_h = P_h^{-1}(X)$ l'hypothèse H8) du théorème 5.1 chap. I est vérifiée. D'où le théorème.

II-3 ETUDE D'UNE EQUATION NON LINEAIRE.

Soit Ω un ouvert borne de $\mathbb{R}^n$ suffisamment régulier. On munit l'espace $W_0^{1,p}(\Omega)$ de la norme $\|u\| = \|\nabla u\|_{L_p(\Omega)}$

avec $\nabla u = \left(\sum_{i=1}^n |D_i u|^2 \right)^{1/2}$ qui en fait un Banach uni-

formément convexe.

On cherche $u \in W_0^{1,p}(\Omega)$ vérifiant

$$(3.1) \quad Au = f \quad \text{pour } f \text{ donné dans } W^{-1,q}(\Omega)$$

$$\text{Avec } Au = - \sum_{i=1}^n D_i \left(|\nabla u|^{p-2} D_i u \right); \quad \nabla u = \left(\sum_{i=1}^n |D_i u|^2 \right)^{1/2}$$

Nous avons la

Proposition 3.1

Pour tout $f \in W^{-1,q}(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$, il existe $u \in W_0^{1,p}(\Omega)$ unique solution de (3.1).

Démonstration

On montre facilement que l'opérateur A est hémiocontinu.

De plus $(Au, u) = \|u\|^p$ et $(Au, v) \leq \|u\|^{p-1} \|v\|$

Donc $(Au - Av, u - v) \geq (\|u\|^{p-1} - \|v\|^{p-1})(\|u\| - \|v\|)$

On applique alors la remarque 1.1 du chap. I.

On se propose maintenant d'approcher la solution $u \in V$ de (3.1)
Pour plus de commodité pour le lecteur, nous reprenons les notations habituelles.

Soit $\mathcal{R}_h$ le réseau des points $M = (m_1, h_1, \dots, m_n, h_n)$ $m_i \in \mathbb{Z}$

$i = 1, \dots, n$. Nous posons :

$$\omega_{n,q}^H = \prod_{i=1}^n \left[\left(m_i - \frac{q_i + 1}{2} \right) h_i, \left(m_i + \frac{q_i + 1}{2} \right) h_i \right] \quad q = (q_i) \quad 1 \leq i \leq n$$

où q_i est un entier positif et $|q| = \sum_{i=1}^n q_i$.

$$\rho_h^H = \bigcup_{|q| \leq 1} \omega_{h,q}^H ; \quad \Omega_h^H = \{M \mid M \in \mathcal{R}_h, \rho_h^H \subset \Omega\} ; \quad \Omega_h = \bigcup_{M \in \Omega_h^H} \omega_{h,0}$$

$\theta_h^H =$ fonction caractéristique de $\omega_{h,0}^H$

$V_h =$ espace des suites de nombres réels $u_h = (u_h^M)_{M \in \Omega_h^H}$

C'est un espace de dimension finie. Si $u_h \in V_h$; u_h est identi-

fié à la fonction $u_h = \sum_{M \in \Omega_h^H} u_h^M \theta_h^M$ Posons

$$\nabla_i u_h(x) = \frac{u_h(x + \frac{h_i}{2}) - u_h(x - \frac{h_i}{2})}{h_i} ; \quad \nabla_i^2 u_h = \nabla_i (\nabla_i u_h)$$

On munit alors V_h de la norme $\|u_h\|_h = \|\nabla_h u_h\|_{L_p}$

$$\text{où } \nabla_h u_h = \left(\sum_{i=1}^n |\nabla_i u_h|^2 \right)^{1/2}.$$

Soit $W_i^{1,p}(\Omega) = \{ \varphi \mid \varphi \in L_p(\Omega), D_i \varphi \in L_p(\Omega) \}$

On sait cf [1] qu'il existe $P_h^i \in \mathcal{L}(V_h, W_i^{1,p}(\Omega))$ et

$r_h \in \mathcal{L}(W_i^{1,p}(\Omega), V_h)$ tels que la famille

$\{ W_i^{1,p}(\Omega), V_h, P_h^i, r_h \}$ soit stable et consistante

et l'on a : $D_i P_h^i u_h = \nabla_i u_h$

A l'équation (3.1) nous associons alors l'équation

$$(3.2) \quad A_h u_h = f_h$$

$$\text{avec (3.3)} \quad A_h u_h = - \sum_{i=1}^n \nabla_i (|\nabla_h u_h|^{p-2} \nabla_i u_h)$$

$$(3.4) \quad f_h = \frac{1}{h_1 \dots h_n} \int_{\Omega} f \theta_h^N dx$$

Nous avons alors la

Proposition 3.2

L'équation (3.2) admet une solution unique $u_h \in V_h$.

De plus $p_h^i u_h \longrightarrow u$ dans $W^{1,p}_i(\Omega)$ fort $|i| \leq 1$; u étant la solution de (3.1).

Dans [22] nous donnons une méthode itérative permettant de résoudre numériquement l'équation (3.2), (3.3), (3.4).

II-4 ETUDE D'UN PROBLEME DE MECANIQUE.

Soit à résoudre l'inéquation

$$(4.1) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X \text{ avec}$$

$$(4.2) \quad X = \left\{ v \in H^1_0(\Omega) ; |\text{grad } u| \leq 1 \text{ p.p sur } \Omega \right\}$$

Ω étant un ouvert borné de R^n .

$$\text{On prendra } A = -\Delta = - \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} + \lambda \text{ Id} \quad \lambda \geq 0$$

et f donné tel que $\lambda \geq |\text{grad } f|$

L'inégalité (4.1) est le problème physique correspondant à la torsion élastoplastique d'une barre (cf. [7], [8], [13] et [24]).

On se propose maintenant de résoudre le problème (4.1) à l'aide de la méthode de la mire du paragraphe I-8.

Auparavant, nous avons le

Théorème 4.1 (cf. H. Brezis et M. Sibony [7])

1 - Le problème (4.1) équivaut à trouver $\tilde{u} \in \tilde{X}$ solution de

$$(4.3) \quad (A \tilde{u}, \tilde{v} - \tilde{u}) \geq (f, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X} \quad \text{avec}$$

$$(4.4) \quad \tilde{X} = \{ \tilde{v} \in H_0^1(\Omega) \mid -\delta \leq v \leq \delta \text{ sur } \Omega \}$$

$$A = -\Delta + \lambda \quad \text{avec } \lambda \geq |\text{grad } f|$$

$$\delta(x) = \text{distance}(x, \partial\Omega)$$

2 - Les problèmes (4.1) et (4.3) admettent une solution unique et on a $u = \tilde{u}$

3 - La solution u de (4.1) est alors donnée par

$$(4.5) \quad u = P_{\tilde{X}}^A (A^{-1} f)$$

Le calcul numérique de $\tilde{u}$ solution de (4.3) peut se faire de plusieurs manières.

1 - On calcule $v = A^{-1} f$ par la méthode de Gauss-Seidel [25] et on projette v sur $\tilde{X}$ suivant la norme $[u] = (Au, u)^{1/2} = \left(\sum_{i=1}^n \|D_i u\|_{L^2}^2 \right)^{1/2}$ à l'aide de la formule (4.5).

2 - On applique la méthode de la mire à la fonctionnelle

$$(4.6) \quad F(\tilde{v}) = \frac{1}{2} \sum_{i=1}^n \|D_i \tilde{v}\|_{L^2}^2 - (f, \tilde{v})$$

avec les contraintes $\tilde{v} \in \tilde{X} = \{ \tilde{v} \in H_0^1(\Omega) \mid -\delta \leq \tilde{v} \leq \delta \text{ sur } \Omega \}$

La méthode basée sur la projection sur $\hat{X}$ est exposée dans [22] où nous donnons aussi des résultats numériques.

Nous étudions ici la résolution numérique de l'inégalité (4.3) à l'aide de la méthode de la mire. D'après le théorème 2.1 chap. I; cela revient à minimiser la fonctionnelle (4.6) sur le convexe

$$(4.7) \quad \tilde{X} = \{ \tilde{v} \in H_0^1(\Omega) ; |\tilde{v}| \leq \delta \text{ p.p sur } \Omega \} ,$$

Les expériences numériques seront faites pour un ouvert $\Omega = (0,1) \times (0,1) \subset \mathbb{R}^2$.

On note par $\{u_{i,j}\}$ les valeurs de u aux points $\{x_{i,j}\}$ $i = 0, \dots, n+1$, $j = 0, \dots, n+1$; qui constituent l'ensemble des points du réseau R_h associé au pas h de discrétisation ($h = 1/(n+1)$).

La relation (4.6) s'écrit alors en dimension finie :

$$(4.8) \quad F(\{u_{i,j}\}) = \frac{1}{4} \sum_{\substack{i=1,n \\ j=1,n}} (u_{i+1,j} - u_{i-1,j})^2 + (u_{i,j+1} - u_{i,j-1})^2 \\ + \frac{1}{2} \sum_{j=1,n} (u_{1,j}^2 + u_{n,j}^2) + \frac{1}{2} \sum_{i=1,n} (u_{i,1}^2 + u_{i,n}^2) \\ - 2h^2 \sum_{\substack{i=1,n \\ j=1,n}} f_{i,j} u_{i,j}$$

Les conditions aux limites s'écrivent en dimension finie :

$$(4.9) \quad u_{i,j} = 0 \quad \text{pour} \quad i = 0, \quad i = n+1 \quad \text{et} \quad j = 0, \quad j = n+1.$$

Donc $F(\{u_{i,j}\})$ est une fonction de n^2 variables $\{u_{i,j}\}_{i,j=1,n}$ qui sont les valeurs de u aux limites i,j internes du réseau.

1°) Première expérience : cas où f est quelconque.

On applique directement la méthode de la mire pour minimiser $F(\{u_{i,j}\})$ donné par (4.8) avec $f = f(x,y)$ ($= e^{j(x,y)}$ pour l'expérience) sur le convexe X donné par (4.2) pour un pas de discrétisation $h = 1/10$ (domaine Ω à 100 points intérieurs). Le problème est résolu en plus de

15 mn sur ordinateur UNIVAC 1108. Nous avons pris comme valeur de départ $u_{ij}^0 = 0 \quad \forall i, j$. Le tableau n° 4.1 donne la distribution de $|\text{grad } u(i, j)|$.

Cette première expérience correspond à l'application brutale de la méthode. Les améliorations que nous allons décrire peuvent évidemment être appliquées dans le cadre général.

2°) Deuxième expérience : cas où $f = c^{te}$

On applique la même méthode pour $f = c^{te}$ (= 10 pour l'expérience) pour un pas $h = 1/10$.

Le tableau n° 4.2 donne la répartition de $|\text{grad } u(i, j)|$ aux différents points du réseau $\mathcal{R}_h$ associé à l'ouvert Ω . Nous comparerons dans la suite les résultats du tableau n° 4.2 à ceux du tableau n° 4.4 résolvant le même problème ($f = 10$) après amélioration de la méthode dans la 3ème expérience.

Le temps d'exécution en UNIVAC 1108 est de l'ordre de 15 mn.

$$h = 1/10 \quad f = e^3(x+y)$$

Tableau n° 4.1

.	1	1	1	1	1	1	1	1	1	.
1	.	1	1	1	1	1	1	1	.	1
1	1	.	1	1	1	1	1	.	1	1
1	1	1	.	1	1	1	.	1	1	1
1	1	1	1	.	.	.	1	1	1	1
1	1	1	1	.	.	.	1	1	1	1
1	1	1	1	.	.	.	1	1	1	1
1	1	1	.	1	1	1	.	1	1	1
1	1	.	1	1	1	1	1	.	1	1
1	.	1	1	1	1	1	1	1	.	1
.	1	1	1	1	1	1	1	1	1	.

. signifie $|\text{grad } u_{i,j}| < 1$ au point (i,j)

1 signifie $|\text{grad } u_{i,j}| = 1$ au point (i,j)

Arrêt des itérations quand $\rho = 7,6 \cdot 10^{-6}$

ρ étant le paramètre de cheminement de la méthode de la mire.

Tableau n° 4.2

$$h = 1/10 \quad ; \quad f = c^{te} = 10$$

.	.	☆	☆	☆	☆	☆	☆	☆	.	.
.	.	.	☆	☆	☆	☆	☆	.	.	.
☆	.	.	.	☆	☆	☆	.	.	.	☆
☆	☆	.	.	.	.	.	.	.	☆	☆
☆	☆	☆	.	.	.	.	.	☆	☆	☆
☆	☆	☆	.	.	.	.	.	☆	☆	☆
☆	☆	☆	.	.	.	.	.	☆	☆	☆
☆	☆	.	.	.	.	.	.	.	☆	☆
☆	.	.	.	☆	☆	☆	.	.	.	☆
.	.	.	☆	☆	☆	☆	☆	.	.	.
.	.	☆	☆	☆	☆	☆	☆	☆	.	.

☆ signifie $|grad u| = 1$ au point (i,j)

.

signifie $|grad u| < 1$ au point (i,j)

résultats quand $\rho < 10^{-5}$ ρ étant le paramètre de cheminement de la méthode de la mire.

3°) Amélioration du temps d'exécution

Nous pouvons rendre la méthode de la mire très performante (en ce qui concerne le temps d'exécution), en opérant comme suit :

Au lieu de calculer $F(u_{11}, \dots, u_{kl} + t, \dots, u_{nn})$ aux extrémités d'une mire $M_{i,j,p}$ (de centre (i, j) et de rayon p) à l'aide d'une formule analogue à (4.8), nous calculons cette valeur à l'aide d'une formule approchée utilisant la valeur $F(\{u_{ij}\})$ déjà connue.

Plus précisément nous avons :

$$(4.10) \quad F(u_{11}, \dots, u_{kl} + t, \dots, u_{nn}) = F(\{u_{ij}\}) + M_{kl} + N_{kl} - 2h^2 f_{kl} u_{kl}$$

où M_{kl} et N_{kl} sont donnés par

$$M_{kl} = \begin{cases} \frac{t}{2} (2u_{kl} - u_{k-2,l} - u_{k+2,l} + t) & \text{si } 2 < k < n-1 \quad \forall l \\ \frac{t}{2} (2u_{kl} - u_{k+2,l} + t) & \text{si } k = 2 \quad \forall l \\ \frac{t}{2} (2u_{kl} - u_{k-2,l} + t) & \text{si } k = n-1 \quad \forall l \\ \frac{t}{2} (3u_{kl} - u_{k+2,l} + \frac{3}{2}t) & \text{si } k = 1 \quad \forall l \\ \frac{t}{2} (3u_{kl} - u_{k-2,l} + \frac{3}{2}t) & \text{si } k = n \quad \forall l \end{cases}$$

$$N_{kl} = \begin{cases} \frac{t}{2} (2u_{kl} - u_{k,l-2} - u_{k,l+2} - t) & \text{si } 2 < l < n \quad \forall k \\ \frac{t}{2} (2u_{kl} - u_{k,l+2} + t) & \text{si } l = 2 \quad \forall k \\ \frac{t}{2} (2u_{kl} - u_{k,l-2} + t) & \text{si } l = n-1 \quad \forall k \\ \frac{t}{2} (3u_{kl} - u_{k,l+2} + \frac{3}{2}t) & \text{si } l = 1 \quad \forall k \\ \frac{t}{2} (3u_{kl} - u_{k,l-2} + \frac{3}{2}t) & \text{si } l = n \quad k \end{cases}$$

Les formules (4.10) permettent donc de calculer $F(\{u_i\})$ de proche en proche, sachant qu'au départ $F(0, \dots, 0) = 0$ (cas où la valeur initiale $u^0 = 0$)

Avantages des formules (4.10)

Le calcul de F par la formule ((4.8) comporte :

$3n^2 + 4n + 4$ multiplications

6 divisions

$5n^2$ additions

Alors que le calcul par la formule (4.10) compte au plus :

10 multiplications

4 divisions

8 additions

Dans le cas où $h = 1/10$ ($n = 9$) le calcul par approximation est environ 50 fois plus rapide que le calcul à l'aide des formules (4.8). Une analyse détaillée des formules (4.10) permet de conclure que l'effet cumulatif des erreurs d'approximation est négligeable vis à vis de l'erreur de discrétisation du problème (4.3), (4.4).

Modulo l'introduction de formules d'approximation telles que (4.10) la méthode de la mire semble être une des méthodes directes les plus performantes.

4) Troisième expérience : cas où $f = c^{te}$.

Si $f = c^{te}$ on a vu que le problème (4.1), (4.2) équivaut à résoudre

$$(4.11) \quad (A\tilde{u}, \tilde{v} - \tilde{u}) \geq (f, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X} \quad \text{avec}$$

$$(4.12) \quad A\tilde{u} = -\Delta \tilde{u}$$

$$(4.13) \quad \tilde{X} = \{v \in H_0^1(\Omega) \mid |\tilde{v}| \leq \delta \text{ sur } \Omega\}$$

avec $\delta(x) = \text{distance}(x, \partial\Omega)$

Le problème consiste donc à minimiser la fonctionnelle

$$(4.14) \quad F(v) = \int_{\Omega} |\text{grad } v|^2 dx - 2 \int_{\Omega} f v dx \quad f = c^{te}$$

sur $\tilde{X}$ donné par (4.13) et non sur X donné par (4.2).

Le tableau n° 4.3 donne la répartition des contraintes pour $h = 10^{-1}$ (réseau à 121 points) et $f = c^{te} = 5$

Le tableau n° 4.4 donne les mêmes résultats pour $f = c^{te} = 10$.

Le temps d'exécution de chacune de ces deux expériences ($f = 5$ et $f = 10$) jusqu'à $\rho \leq 10^{-5}$ est de 1mm en C.D.C. 3.600.

Par ailleurs la comparaison des résultats complets pour u et $|\text{grad } u|$ dans les expériences 2) et 4) conduisent à penser que les résultats de 4) évitent beaucoup les erreurs d'arrondi et sont nettement plus précis que ceux de 2). En effet dans 2) le calcul de $|\text{grad } u|$ (au lieu de $|u|$ dans 4) à chaque itération, introduit beaucoup d'erreurs et un temps d'exécution non négligeable.

Tableau n° 4.3

$$f = c^{te} = 5 ; \quad h = 10^{-1}$$

Tableau de la valeur du gradient

.	.	1	1	1	1	1	1	1	.	.
.	.	.	1	1	1	1	1	.	.	.
1	.	.	.	.	.	.	.	.	.	1
1	1	.	.	.	.	.	.	.	1	1
1	1	.	.	.	.	.	.	.	1	1
1	1	.	.	.	.	.	.	.	1	1
1	1	.	.	.	.	.	.	.	1	1
1	1	.	.	.	.	.	.	.	1	1
1	.	.	.	.	.	.	.	.	.	1
.	.	.	1	1	1	1	1	.	.	.
.	.	1	1	1	1	1	1	1	.	.

. signifie $|\text{grad } u| < 1$ au point (i,j)

1 signifie $|\text{grad } u| = 1$ au point (i,j)

Arrêt des calculs quand $\rho = 3.10^{-3}$

En poursuivant les itérations jusqu'à $\rho = 3.10^{-6}$ on obtient approximativement les mêmes résultats pour $|\text{grad } u|$ et $u(i,j)$.

Tableau n° 4.4

$$h = 1/10 \quad ; \quad f = c^{te} = 10$$

.	.	☆	☆	☆	☆	☆	☆	☆	.	.
.	.	.	☆	☆	☆	☆	☆	.	.	.
☆	.	.	.	☆	☆	☆	.	.	.	☆
☆	☆	.	.	.	.	.	.	.	☆	☆
☆	☆	☆	.	.	.	.	.	☆	☆	☆
☆	☆	☆	.	.	.	.	.	☆	☆	☆
☆	☆	☆	.	.	.	.	.	☆	☆	☆
☆	☆	.	.	.	.	.	.	.	☆	☆
☆	.	.	.	☆	☆	☆	.	.	.	☆
.	.	.	☆	☆	☆	☆	☆	.	.	.
.	.	☆	☆	☆	☆	☆	☆	☆	.	.

☆ signifie $|grad u| = 1$ au point (i,j)

.

signifie $|grad u| < 1$ au point (i,j)

Le tableau 4.5 donne la répartition du $|grad u|$ pour $f = 15$

La comparaison des 3 tableaux: 4.3, 4.4, 4.5 rend compte de la répartition des contraintes $X = \{u \mid |grad u| \leq 1 \text{ p.p}\}$

lorsque $f = c_i^{te}$ ($c_1 = 5$, $c_2 = 10$, $c_3 = 15$) croît.

Tableau n° 4.5

Tableau de la valeur du gradient

Arrêt des calculs quand $\rho = 3.10^{-3}$ les résultats sont approximativement les mêmes si on poursuit les itérations jusqu'à $\rho = 9.10^{-5}$.

.	1	1	1	1	1	1	1	1	1	.
1	.	1	1	1	1	1	1	1	.	1
1	1	.	1	1	1	1	1	.	1	1
1	1	1	.	1	1	1	.	1	1	1
1	1	1	1	.	.	.	1	1	1	1
1	1	1	1	.	.	.	1	1	1	1
1	1	1	1	.	.	.	1	1	1	1
1	1	1	.	1	1	1	.	1	1	1
1	1	.	1	1	1	1	1	.	1	1
1	.	1	1	1	1	1	1	1	.	1
.	1	1	1	1	1	1	1	1	1	.

(.) signifie $|\text{grad } u_{ij}| < 1$ au point (i,j)

(1) signifie $|\text{grad } u_{ij}| = 1$ au point (i,j)

5) Conclusion sur la méthode de la mire.

La méthode de la mire semble très performante moyennant les améliorations du point 3). On peut aussi mener les calculs de la manière suivante :

on se donne les paramètres $h_1 = h_2 = h = 1/5$; on calcule $u_{h,p}$ à l'aide de la méthode de la mire pour ρ suffisamment petit, en prenant comme valeur initiale à la première itération $u_{h,p}^0 = 0$.

Puis on fait $h_1 = h_2 = h/2$ et $u_{\frac{h}{2},p}^0 = \tilde{u}_{h,p}$ où $\tilde{u}_{h,p}$ est une interpolation de $u_{h,p}$ pour le réseau de pas $h_1 = h_2 = h/2$. On peut ensuite reproduire le procédé pour $h_1 = h_2 = h/4$.

Cette méthode donne de bons résultats (nous l'avons appliquée à d'autres problèmes cf [10]).

Il est facile de voir que la méthode de la mire est d'autant plus rapide que le convexe X dans lequel on cherche la solution est "petit". Nous l'appliquons aussi pour minimiser une fonctionnelle dans l'espace entier cf les problèmes du § II-5. Néanmoins, elle devient onéreuse quand le nombre de points de discrétisation augmente beaucoup, surtout en dimension $n = 2$.

II-5 APPLICATIONS DE LA METHODE DE LA MIRE.

Problème 1

On cherche à minimiser dans $H_0^1(\Omega)$, $\Omega = (0,1) \times (0,1)$, la fonctionnelle

$$(5.1) \quad J(v) = \sum_{i=1}^2 \|D_i v\|_{L_2}^2 - 2(f, v)$$

ce qui correspond aussi à résoudre dans $H_0^1(\Omega)$ l'équation

$$(5.2) \quad -\Delta u = f$$

Ce problème est un cas particulier du précédent (avec contraintes) mais c'est un bon moyen de contrôler la précision de la méthode de la mire.

On se donne

$$(5.3) \quad u = (x - x^2)(y - y^2)$$

$$\text{et } (5.4) \quad f = -\Delta u = 2(x - x^2 + y - y^2)$$

On pose alors

$$(5.5) \quad f_h = f \quad \text{dans}$$

$$(5.6) \quad J_h(v_h) = \sum_{i=1}^2 \|\nabla_i v_h\|_{L^2}^2 - 2(f_h, v_h)$$

et on minimise $J_h(v_h)$ dans l'espace

$$V_h = \{ v_h \mid v_h = \sum_M \theta_h^M v_h^M \quad v_h|_\Gamma = 0 \}$$

à l'aide de la méthode de la mire (cf chap. I paragraphe I-8),

en prenant comme valeur initiale $u_h^0 = 0$. Nous comparons alors

le résultat obtenu $u_{h,p}$ ($p < 10^{-5}$) à u donné par (5.3).

La méthode de la mire brute (cf première expérience du problème précédent) donne pour $h = 1/8$:

$$\left\| \frac{u_{h,p} - u}{u} \right\|_{L^\infty} \leq 4.10^{-5}, \quad u \text{ étant la solution exacte donnée par (5.3).}$$

Je remercie J.P. Yvon qui a programmé ce problème.

Problème 2

On cherche $u \in H_0^1(0,1)$ solution de

$$(5.7) \quad \inf_{u \in H_0^1(\Omega)} J(u)$$

$$\text{avec} \quad J(u) = \int_0^1 (u'^2 + u' + 2u) dx$$

L'équation d'Euler donne une solution exacte à ce problème :

$$(5.8) \quad u(x) = -1 + \frac{1}{1+e} (e^x + e^{1-x})$$

On discrétise le problème avec un pas $h = 1/20$ et on applique la méthode de la mire en prenant comme valeur initiale

$$u^0(i) = 0 \quad i = 1, \dots, N+1$$

Au bout de 360 itérations, nous obtenons les résultats suivants :

x	u	u_h	$\left \frac{u_h - u}{u} \right $
0	0	0	
0.05	- 0.021865	- 0.022827	5 %
0.10	- 0.41285	- 0.041016	0.7 %
0.15	- 0.058307	- 0.058838	0.9 %
0.20	- 0.072974	- 0.072510	0.7 %
0.25	- 0.085324	- 0.085571	0.3 %
0.30	- 0.095386	- 0.094604	0.8 %
0.35	- 0.10319	- 0.10315	0.04 %
0.40	- 0.10874	- 0.10779	0.95 %
0.45	- 0.11207	- 0.111820	0.25 %
0.50	- 0.11318	- 0.112060	1 %

Problème 3

On cherche $u \in H_0^1(0,1)$ solution de

$$(5.9) \quad J(u) \leq J(v) \quad \forall v \in H_0^1(0,1)$$

avec
$$J(u) = \int_0^1 (u'^2 + 2e^u) dx$$

L'équation d'Euler donne pour solution

$$(5.10) \quad u(x) = 2 \operatorname{Log} \lambda - \operatorname{Log} 2 - 2 \operatorname{Log} \left(\cos \frac{\lambda}{2} \left(x - \frac{1}{2} \right) \right)$$

où λ est la solution positive de l'équation

$$(5.11) \quad \lambda = \sqrt{2} \cos \frac{\lambda}{4}$$

ce qui donne $\lambda = 1,3361$

On applique alors la méthode de la mire au problème (5.9) avec un pas h discrétisation $h = 1/20$; ce qui donne une solution u_h que nous comparons à u donné par (5.10).

Nous obtenons les résultats suivants au bout de 301 itérations :

x	u	u_h	$\left \frac{u - u_h}{u} \right $
0	0	0	
0.05	- 0.0 22461	- 0.021943	2.5 %
0.10	- 0.0 4064	- 0.041438	2 %
0.15	- 0.0 58105	- 0.058533	0.7 %
0.20	- 0.0 72021	- 0.073270	1.7 %
0.25	- 0.0 84473	- 0.085683	1.4 %
0.30	- 0.0 94238	- 0.095802	1.7 %
0.35	- 0.1 0205	- 0.10365	1.6 %
0.40	- 0.1 0742	- 0.10924	1.6 %
0.45	- 0.1 1084	- 0.11259	1.4 %
0.50	- 0.1 1182	- 0.11371	1.7 %

(Nous avons pris comme vecteur initial $u^0 = 0$)

Je remercie Glowinski pour avoir programmé ce problème.

B I B L I O G R A P H I E

[1] Aubin J.P.

Approximation des espaces de distributions et des opérateurs différentiels. Bull. Société Math. France memoire 12, 1967, 139 p.

[2] Aubin J.P.

Behaviour of the error of the approximate solutions of boundary value problems for linear elliptic operators by Galerkin's and finite difference methods. Annali della Scuola Norm. Pisa (3) 21 (1967), 599-637

[2'] Banitchouk N.V., Petrov V.M., Tchernoussko F.L.

Résolution numérique de problèmes aux limites variationnels par la méthode des variations locales. Journal de Calcul Numérique et de Physique Mathématique, tome 6, n° 6, Moscou 1966. p.947-961

[3] Brezis H.

Les opérateurs monotones, Séminaire Choquet (Initiation à l'Analyse) 5ème année 1965/66 n° 10

[4] Brezis H.

Equations et Inéquations non linéaires dans les espaces vectoriels en dualité. Annales de l'Institut Fourier de l'Université de Grenoble, Tome XVIII Fasc 1, 1968

[5] Brezis H.

Sur la convergence forte de la méthode d'éclatement pour les opérateurs non linéaires (à paraître)

[6] Brezis H. et Sibony M.

Méthodes d'approximation et d'itération pour les opérateurs monotones. Archive for Rat. Meca. and Analysis Vol 28 n° 1 1968 p. 59-82

[7] Brezis H. et Sibony M.

Sur l'équivalence de certaines inéquations variationnelles. (A paraître dans Archive for Rat. Meca and Analysis)

[8] Brezis H. et G. Stampacchia

Sur la régularité de la solution d'inéquations elliptiques. Bull. Soc. Math. France 96, 1968 p. 153-180

[9] Browder F.

Existence and uniqueness theorems for solutions of non linear boundary value problems. Application of non linear partial differential equations in mathematical physics. p. 24-49. Providence Amer. Math. Soc. 1965 (Proceeding of symposia in applied mathematics 17)

[10] Browder F.

Problèmes non linéaires. Université de Montréal 1966

[11] Céa J.

Approximation variationnelle des problèmes aux limites. Annales Inst. Fourier 14 (1964) p. 345-444

[11'] Cherruault Y., G. Nissen, F. Tonnellier

Etude théorique et pratique d'une méthode directe de minimisation et applications. Rapport contrat D.G.R.S.T., n° 6701047. Centre National de la Recherche Scientifique. Publication n° MM/C/9.9.8/AI Oct. 1968

[12] Courant R.

Variational methods for the solution of problems of equilibrium and vibration (Bull A M S 49 p. 1-23) 1943

[13] Hartman P. and Stampacchia G.

On some non linear elliptic differential functional equations. Acta Math. t. 115 (1966) p. 271-310

[14] Leray J. et Lions J.L.

Quelques résultats de Visik sur les problèmes elliptiques non linéaires par les méthodes de Minty-Browder. Bull. Soc. Math. France t. 93 (1965) p. 97-107

[15] Lions J.L.

Problèmes aux limites dans les équations aux dérivées partielles. Université de Montréal 1962

[16] Lions J.L.

Contrôle des systèmes gouvernés par des équations aux dérivées partielles (Dunod 1968)

[17] Lions J.L. et Stampacchia G.

Variational inequalities. Comm. Pure and Applied Math. Sept 1967 t. 20, p. 493-519

- [18] Lions J.L. et Temam R.
Une méthode d'éclatement des opérateurs et des contraintes en calcul des variations. C.R. Acad. Sc. Paris 263, série A (1966) n. 563-565
- [19] Mosco U.
A remark on a theorem of F.E. Browder. Journal of Math. Analysis and Applications vol. 20 n° 1 Oct. 1967
- [20] Sibony M.
Minimisation de fonctionnelles non différentiables. (A paraître dans Israël Journal of Mathematics)
- [21] Sibony M.
Contrôle des systèmes gouvernés par des équations aux dérivées partielles (A paraître dans Rendiconti del Seminario Matematico dell'Università di Padova, tome 43)
- [22] Sibony M.
Méthodes itératives pour les équations et inéquations aux dérivées partielles non linéaires de type monotone. (A paraître dans Calcolo)
- [23] Stampacchia G.
Formes bilinéaires coercitives sur les ensembles convexes. C.R. Acad. Sc. Paris t. 258, p. 4413-4416
- [24] Tink
Elastic Plastic Torsion. Arch. Rat. Mech. Anal. t. 25 (1967)
- [24'] Vainberg
Variational methods for the study of non linear operator (Holden Day)
- [25] Varga R.S.
Matrix iterative analysis. Englewood Cliffs (Prentice Hall Series in automatic computation). 1962

DEUXIEME PARTIE

METHODES ITERATIVES POUR LES EQUATIONS
ET INEQUATIONS AUX DERIVEES PARTIELLES
NON LINEAIRES DE TYPE MONOTONE

INTRODUCTION

Soit V un espace de Banach réel réflexif, A un opérateur non nécessairement linéaire de V dans son dual V' . On se donne un élément fixe $f \in V'$ et on cherche à résoudre dans V l'équation

$$(1) \quad A u = f$$

On se donne maintenant un convexe fermé X de V . On désigne par $(\cdot, \cdot)$ la dualité entre V et V' .

Plus généralement on cherche $u \in X$ solution du problème

$$(2) \quad (A u, v - u) \geq (f, v - u) \quad v \in X.$$

Si $X = V$ on retrouve l'équation (1) (cf [3], [14]).

Par ailleurs, minimiser une fonctionnelle F sur le convexe X est un problème qui peut se ramener à résoudre une inégalité de la forme (2) (cf [11] et [14]).

Des problèmes du genre (1) et (2) interviennent en théorie du contrôle optimal et contiennent une classe assez large d'équation et inéquations aux dérivées partielles linéaires et non linéaires.

Dans [14] nous assurons l'existence et l'unicité de $u \in X$ solution de (2) et nous élargissons le cadre de l'inéquation (2) aux problèmes d'optimisation (cf aussi [13]).

Par ailleurs nous donnons toujours dans [14] de nombreux procédés systématiques pour l'approximation de u solution de (1) et (2).

Ensuite nous approchons l'espace V par des espaces de dimension finie V_h ; ce qui revient à discrétiser le problème.

Il reste alors à donner des méthodes constructives permettant de résoudre le nouveau problème en dimension finie.

On utilise pour cela un procédé itératif valable d'ailleurs en dimension infinie. Il est basé sur le calcul de P_X (projection sur le convexe fermé X).

Nous distinguons alors deux cas :

1er cas

L'opérateur P_X est accessible. Dans ce cas le calcul optimal de certains paramètres, combiné éventuellement avec un changement de norme permet d'aboutir à la solution numérique au bout de quelques itérations seulement.

2ème cas

L'opérateur P_X est accessible.

Nous introduisons alors plusieurs méthodes.

- 1 - On peut supprimer les contraintes X par "pénalisation" ; adaptation d'une idée de Courant (cf. théorème 6.1 de [14] et bibliographie).
 - 2 - On peut "éclater" le convexe $X = \bigcap_{i=1}^n X_i$ en plusieurs X_i en adaptant une méthode de J.L. Lions et R. Temam (cf. théorème 7.1 de [14] et bibliographie).
 - 3 - On "mélange" 1 - et 2 - . (cf. le paragraphe I-7 de [14]).
 - 4 - On peut transformer A et X par changement de variable.
 - 5 - On peut aussi calculer P_X par approximation.
 - 6 - Enfin l'opérateur A est un gradient (1) et (2) équivalent à un problème de minimisation d'une fonctionnelle (cf. [14] § I-2, I-3); problème que l'on peut résoudre par une méthode directe de minimisation, variantes des méthodes des variations locales de Čenauško et Moiseiev et nettement plus opérationnelles.
- Dans le chapitre I nous exposons les méthodes théoriques d'approximation de la solution des problèmes (1) et (2).
- Aux chapitres II et III nous mettons ces idées en application jusqu'au résultats numériques sur de nombreux exemples d'équations d'inéquations aux dérivées partielles et problèmes d'optimisation.

Le plan est le suivant :

I - QUELQUES METHODES GENERALES D'APPROXIMATION.

I-1	<u>Une méthode itérative pour inégalités variationnelles.</u>	
1.1	Itérations avec paramètres fixes.	P.6
1.2	Itérations où le paramètre ρ est variable.	P.8
1.3	Itérations avec changement de paramètre et de norme; accélération de la convergence.	P.13
I-2	<u>Méthode des changements de variables.</u>	P.14
I-3	<u>Méthodes itératives régularisantes.</u>	P.19
3.1	Méthode I : Régularisation et passage à la limite.	P.19
3.2	Méthode II : Méthode itérative régularisante.	P.22
3.3	Méthode III : Méthode mixte.	P.24

II - EQUATION AUX DERIVEES PARTIELLES NON LINEAIRES ET PROBLEME D'OPTIMISATION SUR L'ESPACE ENTIER.

II-1	<u>Problème 1</u>	P.28
	-Résultats numériques du problème 1.	P.33
II-2	<u>Problème 2</u>	P.37
	-Influence du paramètre ϵ .	P.38
	-Influence du paramètre ρ .	P.45
	-Accélération de la convergence.	P.49
II-3	<u>Problème 3</u>	P.49
	1ère méthode : méthode de décomposition de l'opérateur A_h .	P.50
	2ème méthode : méthodes itératives directes.	P.57
II-4	<u>Problème 4</u> : régularisation elliptique.	P.64
II-5	<u>Problème 5</u>	P.67
	-Une nouvelle estimation du paramètre ρ .	P.68
II-6	<u>Problème 6</u>	P.71

III - INEQUATIONS AUX DERIVEES PARTIELLES NON LINEAIRES ET PROBLEME

D'OPTIMISATION SUR UN CONVEXE FERME. P.75

III-1 Problème 1 : Le convexe X est une boule. P.76

-Résultats numériques : un exemple unidimensionnel. P.82

-Résultats numériques : Exemple bidimensionnel. P.85

III-2 Problème 2 : Régularisation elliptique. P.87

III-3 Problème 3 : Le convexe X est un cône. P.89

-1ère expérience numérique. P.95

-2ème expérience numérique. P.97

-3ème expérience numérique; méthode de pénalisation. P.99

III-4 Problème 4 P.105

-Un exemple unidimensionnel : X est un convexe sur lequel on ne sait projeter directement.

-1ère méthode. P.106

-2ème méthode. P.109

-Résultats numériques. P.111

III-5 Un problème unidimensionnel : X est un convexe sur lequel on ne sait pas projeter explicitement. P.115

III-6 Résolution d'un problème de mécanique.

Résultats numériques. P.125

I - QUELQUES METHODES GENERALES D'APPROXIMATION

I - 1. UNE METHODE ITERATIVE POUR INEGALITES VARIATIONNELLES.

Soit $\mathcal{H}$ un Hilbert sur $\mathbb{R}$ muni du produit scalaire $(\cdot, \cdot)$ et de la norme $\|\cdot\|$. Soit X un convexe fermé de $\mathcal{H}$ contenant l'origine. On se propose de résoudre par une méthode itérative l'inéquation:

$$(1.1) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X, \\ \text{pour } f \text{ donné dans } \mathcal{H}$$

où A est un opérateur non nécessairement linéaire de $\mathcal{H}$ dans $\mathcal{H}'$

Dans le cas où l'opérateur A applique un banach V dans son dual, on se ramène après discrétisation à une inégalité de la forme:

$$(1.2) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h \\ \text{pour } f_h \text{ donné dans } V'_h;$$

où X_h est un convexe fermé d'un espace de dimension fini V_h associé au paramètre h de discrétisation. L'espace V_h est muni du produit scalaire $(\cdot, \cdot)_h$ et l'opérateur A_h applique V_h dans son dual V'_h .

Pour effectuer cette discrétisation on se sert du théorème 5.1 de [14] ou d'autres méthodes cf. [6].

L'inégalité (1.2) est alors un cas particulier de (1.1).

Dans de nombreux cas le problème (1.1) est posé dans un hilbert $\mathcal{H}$ sous espace de

$$H^m(\Omega) = \{u \mid u \in L^2(\Omega), D_i u \in L^2(\Omega) \quad |i| \leq m\}$$

Ω étant un ouvert borné de $\mathbb{R}^n$.

Dans ce cas on peut appliquer directement la méthode itérative qui suit sans passer par l'étape de discrétisation (1.2).

La difficulté pratique de notre méthode est dans le calcul numérique de $P_X(u)$ quand $u \notin X$, où P_X est l'opérateur de projection sur le convexe X .

Dans la suite de ce chapitre nous donnons diverses méthodes permettant de résoudre cette difficulté dans de nombreux cas pratiques cf. aussi les paragraphes I-6, I-7 et I-8 de [14]. Dans le paragraphe I-8

en particulier, nous donnons une méthode conduisant au calcul numérique de la projection sur un convexe quelconque de $\mathbb{R}^n$.

1.1. Itérations avec paramètres fixes.

On se donne une application $S \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ vérifiant H1)

H1) i) $(Su, v) = (u, Sv) \quad \forall u, v \in \mathcal{H}$

ii) $\exists$ une constante $\sigma > 0$ telle que $(Su, u) \geq \sigma \|u\|^2$.

Posons $[u]^2 = (Su, u)$; alors $[u]$ définit sur $\mathcal{H}$ une norme équivalente à $\| \cdot \|$.

On suppose que l'opérateur A vérifie l'hypothèse

H2) i) $(Au - Av, u - v) \geq k [u - v]^2 \quad \forall u, v \in \mathcal{H} \quad k > 0$.

ii) $\forall N$ constante positive $\exists$ une constante $C(N) > 0$ telle que $\forall u, v \in \mathcal{H}$ avec $[u] \leq N, [v] \leq N$ alors

$(Au - Av, w) \leq C(N) [u - v] [w] \quad \forall w \in \mathcal{H}$.

Lemme 1.1

Si P_X^S est l'opérateur de projection sur X au sens du produit scalaire $[u, v] = (Su, v)$ alors toute solution de (1.1) est solution de

(1.3) $P_X^S(u - \rho S^{-1}(Au - f)) = u \quad \rho > 0$

et réciproquement.

Démonstration

Posons $g = \rho(Au - f)$ avec $\rho > 0$. Alors (1.3) s'écrit :

$P_X^S(u - S^{-1}g) = u$; ce qui équivaut à $[-S^{-1}g, v - u] \leq 0 \quad \forall v \in X$

c'est-à-dire, $(-g, v - u) \leq 0 \quad \forall v \in X$ ce qui équivaut à $(Au - f, v - u) \geq 0 \quad \forall v \in X$.

D'où le

Théorème 1.1

On fait les hypothèses H1) et H2) alors

1°) $\forall f \in \mathcal{H}' \quad \exists u \in X$ unique solution de

(1.1) $(Au, v - u) \geq (f, v - u) \quad \forall v \in X$

2°) La suite (u_n) définie par l'itération :

(1.4) $u_{n+1} = P_X^S(u_n - \rho S^{-1}(Au_n - f))$ converge fortement dans $\mathcal{H}$ vers u pour $\rho > 0$ et u_0 convenablement choisis ($u_0 = 0$ par exemple).

Démonstration.

1°) L'existence et l'unicité de la solution résultent du théorème 1.1 de [14] .

2°) Posons $\epsilon_{n+1} = u_{n+1} - u$. Par soustraction de (1.4) et (1.3) on a :

$$\begin{aligned}\epsilon_{n+1} &= P_X^S(u_n - \rho S^{-1}(Au_n - f)) - P_X^S(u - \rho S^{-1}(Au - f)) \\ [\epsilon_{n+1}]^2 &= [P_X^S(u_n - \rho S^{-1}(Au_n - f)) - P_X^S(u - \rho S^{-1}(Au - f))]^2\end{aligned}$$

ce qui donne :

$$(1.5) \quad [\epsilon_{n+1}]^2 \leq [\epsilon_n - \rho S^{-1}(Au_n - Au)]^2. \text{ Or,}$$

$$[\epsilon_n - \rho S^{-1}(Au_n - Au)]^2 = (S \epsilon_n, \epsilon_n) - 2\rho (Au_n - Au, \epsilon_n) + \rho^2 (Au_n - Au, S^{-1}(Au_n - Au))$$

Donc

$$[\epsilon_{n+1}]^2 \leq [\epsilon_n]^2 - 2\rho (Au_n - Au, \epsilon_n) + \rho^2 [S^{-1}(Au_n - Au)]^2.$$

Soit N_0 tel que $[u] \leq \frac{N_0}{2}$. On fait l'hypothèse de récurrence :

$$[\epsilon_n] = [u_n - u] \leq \frac{N_0}{2}, \text{ ce qui entraîne, } [u_n] \leq N_0 \text{ et d'après H2),}$$

$$\exists C(N_0) = C \text{ tel que}$$

$$(Au_n - Au, S^{-1}(Au_n - Au)) \leq C[\epsilon_n] [S^{-1}(Au_n - Au)]. \text{ De plus H2) donne}$$

$$(Au_n - Au, \epsilon_n) \geq k[\epsilon_n]^2. \text{ D'où}$$

$$(1.6) \quad [\epsilon_{n+1}]^2 \leq (1 - 2\rho k + \rho^2 C^2)[\epsilon_n]^2 = \theta[\epsilon_n]^2.$$

On choisit $\rho > 0$ tel que $\theta < 1$ et l'hypothèse de récurrence est vérifiée. Par conséquent $\epsilon_n \longrightarrow 0$ quand $n \longrightarrow \infty$, ce qui achève la démonstration.

Remarque 1.1

1°) En particulier la suite (u_n) converge vers u pour $u_0 = 0$

$$N_0 = \frac{2}{k\sqrt{C}} \|f - A(o)\|, \quad C = C(N_0) \text{ et } \rho_{\text{opt}} = k/C^2;$$

$$\text{ce qui donne } \theta = 1 - \frac{k^2}{C^2} < 1.$$

2°) Nous avons alors la majoration a priori de l'erreur :

$$\sqrt{\sigma} \|u_n - u\| \leq [u_n - u] \leq \theta_S \frac{n}{2} [u] \text{ avec } \theta_S = 1 - \frac{k}{C^2(N)}$$

Si on prend $S = \text{Id}$ on a

$$\|u_n - u\| \leq \theta_I \frac{n}{2} \|u\|.$$

On voit alors que la convergence est d'autant plus rapide que le coefficient θ_S est loin de l'unité. Un bon choix de l'opérateur S assure ceci dans la pratique (cf les exemples du chapitre II).

Remarque 1.2

On peut décomposer les itérations (1.4) en deux étapes de la manière suivante :

$$(1.7) \quad u^{n+\frac{1}{2}} = P_X^S (u^n - \rho S^{-1}(Au^n - f))$$

$$(1.8) \quad u^{n+1} = u^n - \frac{\alpha_n^2}{\alpha_{n+\frac{1}{2}} - \alpha_n}$$

avec $\alpha_n = P_X^S(u^n - \rho S^{-1}(Au^n - f)) - u^n$.

C'est la décomposition d'Aitken. Cette méthode accélère très nettement la convergence (cf. chap. II, problème 2 ; (2.22), (2.23)).

Remarque 1.3

Le théorème 1.1 ci-dessus généralise au cas des inégalités non linéaires le théorème 7.1 de [3]. De plus les itérations (1.4) semblent plus opérationnelles que d'autres méthodes [12], telles que celles du gradient et du gradient projeté [8]. Comme on le verra dans les exemples du chapitre II, les itérations (1.4) ont un domaine d'application très large.

Remarque 1.4

Si l'opérateur A monotone hemicontinu ne vérifie pas l'hypothèse H2) i) on s'y ramène moyennant une régularisation elliptique cf. [14] théorème 4.1.

Dans la suite de ce paragraphe nous allons définir des itérations où les paramètres ρ et S sont variables.

1.2. Itérations où le paramètre ρ est variable.

Tout d'abord nous faisons les remarques suivantes :

1°) Dans les itérations (1.4) on peut poser

$$(1.9) \quad \rho = \rho_1 = \frac{(Au_n - Au_{n-1}, u_n - u_{n-1})}{[S^{-1}(Au_n - Au_{n-1})]^2}$$

dans le calcul de u_{n+1} ou de u_{n+1} et u_{n+2} après avoir démarré les

calculs avec un paramètre ρ fixe ou bien après n_0 itérations effectuées avec ρ fixe donné par la remarque 1.1.

2°) Si on fait $X = V$ on peut alors résoudre l'équation

(1.10) $Au = f$ pour f donné dans $\mathcal{H}'$, avec A linéaire à l'aide des itérations suivantes :

$$(1.11) \quad u_{n+1} = u_n - \rho_n (Au_n - f)$$

où ρ_n est un paramètre positif que nous allons déterminer.

Posons $v_n = Au_n - f$; il vient :

$$(1.12) \quad v_{n+1} = v_n - \rho_n Av_n$$

on a alors

$$\|v_{n+1}\|^2 = \|v_n\|^2 - 2 \rho_n (Av_n, v_n) + \rho_n^2 \|Av_n\|^2$$

$$\text{et } \|v_{n+1}\|^2 \text{ est minimum si } \rho_n = \frac{(Av_n, v_n)}{\|Av_n\|^2}$$

Si par exemple A est un opérateur linéaire sur un Hilbert $\mathcal{H}$ et tel que $(Av_n, v_n) \geq \alpha \|v_n\|^2$ et $\|Av_n\| \leq M \|v_n\|$

$$\text{on a alors } \|v_{n+1}\|^2 \leq \left(1 - \frac{\alpha^2}{M^2}\right) \|v_n\|^2$$

ce qui assure la convergence des itérations (1.11).

3°) Sous les hypothèses du théorème 1.1; on sait que la suite (u_n) définie par les itérations

$$(1.13) \quad u_{n+1} = u_n - \rho(Au_n - f)$$

converge vers u solution de

$$(1.14) \quad Au = f$$

pour u_0 et ρ convenables.

La convergence est d'autant plus rapide que le paramètre ρ est proche de l'optimalité.

Dans certains cas le paramètre ρ fixé calculé à partir de majoration à priori sur l'opérateur ($\rho = k/C^2(N)$) conduit à un coefficient

de contraction $\theta = 1 - \frac{k^2}{C^2(N)}$ très voisin de l'unité; il faudrait

alors plusieurs millions d'itérations pour que le procédé converge ou bien la machine n'est pas assez sensible tout simplement à ce coefficient θ très voisin de 1. On opère alors (avec succès !) de la manière suivante :

1) On se donne u_0 et $\rho_1 = 1$ on calcule alors

$$u_1^1 = u_0 - \rho_1(Au_0 - f)$$

Si $\|Au_1^1 - f\| \leq \|Au_0 - f\|$ on fait $\rho_2 < \rho_1$ (par exemple) et on calcule

$$u_1^2 = u_0 - \rho_2(Au_0 - f)$$

Si $\|Au_1^p - f\| \leq \|Au_1^{p-1} - f\|$ on fait $\rho_{p+1} < \rho_p$

et on calcule

$$u_1^{p+1} = u_0 - \rho_{p+1}(Au_0 - f)$$

2) Pour un certain ρ_p on n'a plus d'amélioration. On pose alors

$$u_1 = u_1^p$$

et on reprend le calcul de

$$u_2^p = u_1 - \rho_p(Au_1 - f)$$

ce qui donne $u_2 = u_2^p$

et ainsi de suite jusqu'à $u_n = u_n^p$ pour n assez grand.

4°) Les itérations (1.4) ou (1.13) convergent bien avec un paramètre

$\rho = k/C^2(N)$ fixe et calculé à partir des hypothèses H2) sur l'opérateur A . Mais en surestimant ce paramètre ρ à partir de sa valeur théorique on obtient une convergence nettement plus rapide ; (au lieu de quelques centaines d'itérations, il suffit parfois de quelques dizaines) cf. les résultats numériques du chap. II.

5°) Dans la remarque 1.1, nous avons vu que dans les itérations (1.4) on pouvait prendre : $u_0 = 0$ et

$$(1.15) \quad \rho = \rho_0 = k/C^2(N_0) \quad \text{avec} \quad N_0 = \frac{2}{k\sqrt{\sigma}} \|f - A(0)\|$$

au lieu de poursuivre les itérations (1.4) avec $\rho = \rho_0$ on peut faire

$$(1.16) \quad \rho = \rho_{n+1} = k/C^2(N_{n+1}) \quad \text{avec} \quad N_{n+1} = \|u_n\|$$

ou bien (ce qui est justifié expérimentalement), on peut poser dans (1.4)

$$\rho = \rho_0 \quad \text{jusqu'à } n = n_0 \quad \text{puis poursuivre pour } n > n_0$$

avec $\rho = \rho_n$ donné par (1.16).

6°) Dans la démonstration du théorème 1.1 on a posé

$$(1.17) \quad N_0 = 2 \|u\| \quad u \text{ étant la solution exacte du problème (1.1) .}$$

Nous allons donner une estimation de cette constante N_0 dans le cas où $X = V$.

Nous avons successivement :

$$\begin{aligned} & \|u\| \leq \|u - u^n\| + \|u^n\| \\ k \|u - u^n\|^2 & \leq (Au^n - Au, u^n - u) \leq \frac{1}{\sqrt{\sigma}} \|f - Au^n\|^* \|u - u^n\| \end{aligned}$$

Donc

$$(1.18) \quad \|u\| \leq \frac{1}{k\sqrt{\sigma}} \|f - Au^n\|^* + \sqrt{\sigma} \|u^n\|$$

on peut alors poser

$$(1.19) \quad N_n = \frac{2}{k\sqrt{\sigma}} \|f - Au^n\|^* + 2\sqrt{\sigma} \|u^n\|$$

Nous avons alors

$$(1.20) \quad \lim_{n \rightarrow \infty} N_n = N_0 = 2\|u\|$$

Dans la pratique $N_n \sim N_0 = 2\|u\|$ après quelques dizaines d'itérations et pour $n > n_0$ nous itérons avec le meilleur paramètre $\rho = \rho_0$ possible.

Sous les hypothèses H1) et H2) on démontre alors la

Proposition 1.1

La suite définie par les itérations

$$(1.21) \quad u_{n+1} = u_n - \rho_n S^{-1}(Au_n - f)$$

$$\text{avec } \rho_n = k/C^2(N_n) \quad \text{et} \quad N_n = \frac{2}{k\sqrt{\sigma}} \|f - Au_n\| + 2\sqrt{\sigma} \|u_n\|$$

converge vers la solution u de l'équation

$$(1.22) \quad Au = f.$$

7°) Dans le cas où $X = V$ les itérations (1.4) deviennent

$$(1.23) \quad u_{n+1} = u_n - \rho S^{-1}(Au_n - f) = u_n - \rho g_n \quad \text{avec} \\ g_n = S^{-1}(Au_n - f)$$

Il s'agit alors de minimiser l'expression

$$(1.24) \quad \varphi(\rho) = \|u_n - u\|^2 = \|u_{n-1} - u - \rho g_{n-1}\|^2$$

u étant la solution de $Au = f$
ce qui donne

$$(1.25) \quad \rho = \frac{(g_{n-1}, u_n - u)}{\|g_{n-1}\|^2}$$

Nous pouvons alors poser dans (1.23) :

$$(1.26) \quad \rho = \rho_n = \frac{(g_{n-1}, u_n - u_{n-1})}{\|g_{n-1}\|^2}$$

on peut toujours démarrer les itérations (1.23) avec un paramètre ρ fixe.

1.3 Itérations avec changement de paramètre et de norme; accélération de la convergence.

On fait les hypothèses H1) et H2) du paramètre 1.1

Si on pose $S = Id$ on a alors les itérations

$$(1.27) \quad u_{n+1} = P_X (u_n - \rho_{Id} (Au_n - f)) \quad \text{avec} \quad \rho_{Id} = k_{Id}/C_{Id}^2/(N)$$

Si on pose $S \neq Id$ on a alors les itérations

$$(1.28) \quad u_{n+1} = P_X^S (u_n - \rho_S S^{-1} (Au_n - f)) \quad \text{avec} \quad \rho_S = k_S/C_S^2 (N)$$

Les itérations (1.27) impliquent que la projection suivant $(,)$ est accessible.

Les itérations (1.28) impliquent que l'on sache projeter suivant le produit scalaire $[,]$.

Nous avons alors la

Proposition 1.2

Supposons que l'on sache projeter suivant $(,)$ et non suivant $[,]$ et que les itérations (1.28) soient plus rapidement convergentes que les itérations (1.27). Alors, il existe un opérateur S_i et un paramètre ρ_i tels que les itérations

$$(1.29) \quad u_{n+1} = P_X^{S_i} (u_n - \rho_i S_i^{-1} (Au_n - f)) \quad \text{soient possibles.}$$

De plus les itérations (1.29) sont plus rapidement convergentes que (1.27).

Démonstration

Posons dans (1.29) $S_i = S$ et $\rho_i = \rho_S$

Si $f_n = u_n - \rho_S S^{-1} (Au_n - f) \in X$ alors

$$(1.30) \quad u_{n+1} = P_X^S (f_n) = f_n$$

et les itérations (1.29) sont plus rapides que (1.27).

Si $f_n = u_n - \rho_S S^{-1} (Au_n - f) \notin X$

On pose $S_i = Id$ et $\rho_i = \rho_{Id}$ et on est conduit aux itérations (1.27) qui sont accessibles.

Remarque 1.5

On peut aussi faire $S_i = Id$ jusqu'à $n = n_0$ et à partir de ce rang n_0 on aura toujours $f_n \in X$. On poursuit alors avec les itérations (1.29).

Pour les applications cf. théorème 3.2 § III-3.

I - 2. METHODE DES CHANGEMENTS DE VARIABLES.

Le théorème 1.1 donne une méthode itérative

$$(2.1) \quad u_{n+1} = P_X^S(u_n - \rho S^{-1}(Au_n - f))$$

permettant de résoudre numériquement le problème :

$$(2.2) \quad (Au, v-u) \geq (f, v-u) \quad \forall v \in X ;$$

où X est un convexe fermé de $\mathcal{H}$.

En général on ne peut calculer explicitement

$P_X^S(u)$ pour $u \notin X$ que pour un nombre restreint de convexes que nous appelons les "bons convexes".

Nous allons donner deux méthodes permettant de résoudre très simplement le problème (2.2) dans de nombreux cas où X n'est pas un "bon convexe".

Dans la première méthode, on "calcule" P_X par transformation de X en un convexe $\tilde{X}$ sur lequel on sait projeter numériquement.

Dans la deuxième méthode on transforme l'inéquation (2.2) en une inéquation de la forme

$$(\tilde{A} \tilde{u}, \tilde{v} - \tilde{u}) \geq (\tilde{f}, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X}$$

$\tilde{X}$ étant un convexe sur lequel on sait projeter explicitement. La solution $\tilde{u} \in \tilde{X}$ de ce problème peut alors être calculée à l'aide des itérations toujours possibles :

$$\tilde{u}_{n+1} = P_{\tilde{X}}^S(\tilde{u}_n - \tilde{\rho} \tilde{S}^{-1}(\tilde{A}\tilde{u}_n - \tilde{f})).$$

Plus précisément on fait l'hypothèse suivante :

H.3) Il existe un isomorphisme T (pour la structure hilbertienne) de $\mathcal{H}$ dans $\mathcal{H}$ tel que $\tilde{X} = T^{-1}(X)$ soit un "bon convexe".

(une hypothèse plus faible est donnée dans la remarque 2.2)

Nous avons alors le

Théorème 2.1

On se place dans les conditions du théorème 1.1 De plus on suppose H. Posons $f_n = u_n - \rho S^{-1}(Au_n - f)$, n fixé ; alors

1°) La suite $(\tilde{v}_m^n)$ définie par les itérations ;

$$(2.3) \quad \tilde{v}_{m+1}^n = P_{\tilde{X}}^S(\tilde{v}_m^n - \tilde{\rho}(\tilde{B} \tilde{v}_m^n - \tilde{f}_n)), \quad n \text{ fixé}$$

où $\tilde{B} = T^* S T$ et $\tilde{f}_n = T^* S f_n$

converge fortement dans $\mathcal{H}$ vers un élément $\tilde{v}^n$ pour $\tilde{v}_0^n (= 0)$ et $\tilde{\rho} > 0$ convenablement choisis.

2°) La suite (u_n) définie par l'itération

$$(2.4) \quad u_{n+1} = T(\tilde{v}^n)$$

converge fortement dans $\mathcal{H}$ vers u solution de (2.2) pour $u_0 = 0$ et $\rho = k/C^2(N)$.

Démonstration.

Il s'agit de rendre possible les itérations

$$(2.5) \quad u_{n+1} = P_X^E(u_n - \rho S^{-1}(Au_n - f)) = P_X^E(f_n)$$

on part de $u_0 = 0$ par exemple. Or projeter f_n suivant le produit scalaire

$[u, v] = (Su, v)$, équivaut à résoudre :

$$(2.6) \quad (Su_{n+1}, w - u_{n+1}) \geq (Sf_n, w - u_{n+1}) \quad \forall w \in X$$

Posons $u_{n+1} = T\tilde{v}^n$, $w = T\tilde{w}$ pour n fixé.

Il vient alors

$$(S.T\tilde{v}^n, T(\tilde{w} - \tilde{v}^n)) \geq (Sf_n, T(\tilde{w} - \tilde{v}^n)) \quad \forall \tilde{w} \in \tilde{X}$$

ce qui équivaut à

$$(2.7) \quad (\tilde{B}\tilde{v}^n, \tilde{w} - \tilde{v}^n) \geq (\tilde{f}_n, \tilde{w} - \tilde{v}^n) \quad \forall \tilde{w} \in \tilde{X}$$

en posant $\tilde{B} = T^*ST$ et $\tilde{f}_n = T^*Sf_n$.

L'équation (2.7) se résout alors à l'aide des itérations suivantes :

$$(2.8) \quad \tilde{v}_{m+1}^n = P_{\tilde{X}}(\tilde{v}_m^n - \tilde{\rho}(\tilde{B}\tilde{v}_m^n - \tilde{f}_n))$$

qui convergent pour $\tilde{v}_0^n = 0$ et $\tilde{\rho} > 0$ convenablement choisi, vers $\tilde{v}^n$, ne dépendant que de l'indice n .

Les itérations (2.5) deviennent maintenant possibles et s'écrivent :

$$u_{n+1} = T(\tilde{v}^n)$$

D'après le théorème 1.1, on a $u_n \longrightarrow u$ dans $\mathcal{H}$ fort quand $n \longrightarrow \infty$, pour $u_0 (= 0)$ et $\rho = k/C^2(N)$; u étant l'unique solution du problème (2.2).

Remarque 2.1

1°) Si l'opérateur linéaire $\tilde{B} = T^*ST$ vérifie

$$(\tilde{B}\tilde{u}, \tilde{u}) \geq \alpha \|\tilde{u}\|^2 \quad \text{et} \quad (\tilde{B}\tilde{u}, \tilde{v}) \leq c \|\tilde{u}\| \|\tilde{v}\|$$

on peut prendre alors dans (2.8), $\tilde{\rho} = \alpha/C^2$.

On peut aussi faire $\tilde{\rho} = \tilde{\rho}^n$ (cf les remarques du paragraphe I.1, 1.2).

2°) Dans la méthode précédente, pour résoudre l'inéquation

$$(2.9) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X$$

avec A non linéaire et X = convexe fermé sur lequel on ne sait pas projeter explicitement, à l'aide des itérations ;

$$(2.10) \quad u_{n+1} = P_X^S(u_n - \rho S^{-1}(Au_n - f)) ;$$

On se ramène à résoudre à chaque itération n , une inéquation de la forme :

$$(2.11) \quad (\tilde{B} \tilde{u}, \tilde{v} - \tilde{u}) \geq (\tilde{f}, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X}$$

où $\tilde{B}$ est un opérateur linéaire et $\tilde{X}$ est un convexe fermé sur lequel on sait projeter.

Bien entendu on peut résoudre aussi (2.11) avec des itérations analogues à (2.10).

3°) Si $\mathcal{H}' = V_h$ nous avons bien entendu $u_n \longrightarrow u_h$ avec u_h solution d'un problème discret :

$$(A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

et de plus on a $p_h u_h \longrightarrow u$ dans V fort quand $h \longrightarrow 0$ où u est la solution du problème continu :

$$(Au, v - u) \geq (f, v - u) \quad \forall v \in X$$

avec A opérateur cette fois d'un Banach V dans son dual V' qui dans la pratique est un espace de distributions ; cf. [14].

Théorème 2.2

On se place dans les conditions du théorème 1.1. De plus on suppose H3).

Alors

1°) Toute solution u de (2.2) s'écrit

$$(2.12) \quad u = T \tilde{u}$$

où $\tilde{u}$ est solution de l'inégalité

$$(2.13) \quad (\tilde{A} \tilde{u}, \tilde{v} - \tilde{u}) \geq (\tilde{f}, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X} \quad \text{avec} \quad \tilde{A} = T^* A T, \tilde{f} = T^* f, X = T(\tilde{X})$$

et réciproquement

2°) La suite $(\tilde{u}_n)$ définie par les itérations

$$(2.14) \quad \tilde{u}_{n+1} = P_{\tilde{X}}^{\tilde{S}} (\tilde{u}_n - \tilde{\rho} S^{-1} (\tilde{A} \tilde{u}_n - \tilde{f}))$$

converge fortement dans $\mathcal{H}$ vers $\tilde{u}$ solution de (2.13) pour $\tilde{u}_0$ ($= 0$) et $\tilde{\rho} > 0$ convenablement choisis.

Démonstration

1°) Posons dans (2.2) $u = T \tilde{u}$, $v = T \tilde{v}$ il vient alors

$$(2.15) \quad (AT\tilde{u}, T(\tilde{v} - \tilde{u})) \geq (f, T(\tilde{v} - \tilde{u})) \quad \forall \tilde{v} \in \tilde{X}$$

ce qui équivaut à (2.13).

2°) Comme T est un isomorphisme alors l'opérateur $\tilde{A}$ vérifie les hypothèses du théorème 2.1, ce qui assure la convergence de la suite $(\tilde{u}_n)$ vers $\tilde{u}$ solution de (2.13).

Remarque 2.2.

Plus généralement on peut remplacer l'hypothèse H3) par H4)

H4) i) Il existe une application Θ de $\mathcal{H}$ dans $\mathcal{H} \times \mathcal{H}$ linéaire continue et injective telle que $\Theta(X) = \tilde{X}$ soit un "bon convexe".

ii) Il existe une application T de $\mathcal{H} \times \mathcal{H}$ dans $\mathcal{H}$ linéaire continue et surjective telle que $T \Theta = \text{Identité}$.

Nous verrons au paragraphe III-5 un exemple où cette hypothèse est fort utile et simplifie la résolution d'un problème.

Remarque 2.3

1°) Comme on ne suppose pas X borné, notre méthode s'applique au cas où $X = \mathcal{H}$ et permet de résoudre simplement l'équation $Au = f$.

2°) Un choix judicieux des opérateurs S et T peut conduire à un problème

$$(\tilde{A}\tilde{u}, \tilde{v}-\tilde{u}) \geq (\tilde{f}, \tilde{v}-\tilde{u}) \quad \forall \tilde{v} \in \tilde{X}$$

très facile à résoudre. C'est ce que nous verrons dans les exemples du Chapitre II.

3°) Chaque fois que l'opérateur A est linéaire auto-adjoint, en prenant $S = A$ et $[u]^2 = (Au, u)$ l'itération

$$(2.16) \quad u_{n+1} = P_X^A(u_n - \rho S^{-1}(Au_n - f))$$

devient

$$u = P_X^A(A^{-1}f) \quad (\text{car } \rho = 1)$$

ce qui permet de trouver u sans passer par les itérations (2.16).

Cela suppose bien entendu que l'on a une autre méthode plus simple pour inverser l'opérateur $\tilde{A}$ (par exemple si $A = -\Delta$).

Si A est auto-adjoint alors $\tilde{A} = T^*AT$ l'est. On prendra alors $S = \tilde{A}$ et

$[\tilde{u}]^2 = (\tilde{A}\tilde{u}, \tilde{u})$ et l'itération

$$(2.17) \quad \tilde{u}_{n+1} = P_{\tilde{X}}^{\tilde{A}}(\tilde{u}_n - \tilde{\rho} \tilde{S}^{-1}(\tilde{A}\tilde{u}_n - \tilde{f}))$$

devient

$$(2.18) \quad \tilde{u} = P_{\tilde{X}}^{\tilde{A}}(\tilde{A}^{-1}\tilde{f}) = P_{\tilde{X}}^{\tilde{A}}(T^{-1}A^{-1}f). \text{ D'où}$$

$$(2.19) \quad u = T P_{\tilde{X}}^{\tilde{A}}(T^{-1}A^{-1}f)$$

Remarque 2.4

Soit F une fonctionnelle sur un convexe X d'un espace de Hilbert $\mathcal{H}$ sur $\mathbb{R}$. On pose le problème numérique suivant : trouver $u \in X$ tel que

$$(2.20) \quad F(u) \leq F(v) \quad \forall v \in X$$

1°) Si F est différentiable au sens de Gateaux on a vu cf. [14] que toute solution de (2.20) est solution de

$$(2.21) \quad (F'(u), v-u) \geq 0 \quad \forall v \in X$$

et réciproquement.

Pour résoudre numériquement (2.21) on applique alors les méthodes des paragraphes I-1 et I-2.

2°) Si F n'est pas différentiable, on a vu que (2.20) pouvait se mettre sous la forme

$$(2.22) \quad (A_1 u_1, v_1 - u_1) \geq (f_1, v_1 - u_1) \quad \forall v_1 \in X_1 \quad c_f^{[14]}$$

et on applique ensuite les méthodes précédentes pour résoudre (2.22) numériquement.

Dans le chapitre II et dans [14] nous donnons de nombreux exemples d'inéquations variationnelles et de problèmes d'optimisation et les formes (2.20) correspondants.

Remarque 2.5

Pour résoudre numériquement l'inéquation (2.22) il est essentiel de savoir projeter explicitement sur le convexe

$$X_1 = \{v_1 \mid v_1 = (v, \lambda) \in X \times \mathbb{R}, G(v) \leq \lambda\}$$

La proposition qui suit traite un cas particulier important.

Proposition 2.1

Soit $\mathcal{H}$ un Hilbert sur $\mathbb{R}$ de norme $\|\cdot\|$.

On pose

$$(2.23) \quad X_1 = \{v_1 \mid v_1 = (v, \lambda) \in \mathcal{H} \times \mathbb{R}; [v]^2 \leq \lambda\}$$

où $[\cdot]$ est une norme hilbertienne quelconque.

Alors pour tout $v_1 = (v_0, \lambda_0) \notin X_1$ on a :

$$(2.24) \quad P_{X_1}(v_1) = (\theta v_0, \mu)$$

$$(2.25) \quad \mu = \theta^2 [v_0]^2$$

et θ est solution de l'équation :

$$(2.26) \quad \theta^3 + \theta \left(\frac{\|v_0\|^2 - 2\lambda_0 [v_0]^2}{2[v_0]^4} \right) - \frac{\|v_0\|^2}{2[v_0]^4} = 0$$

avec $\theta < 1$ et $\theta^2 [v_0]^2 > \lambda_0$.

Démonstration

Soit $v_1 = (v_0, \lambda_0) \notin X_1$.

Comme la norme $[\cdot]$ est hilbertienne la projection sur X_1 de v_1 est dans le même plan que v_0 . On a donc :

$$P_{X_1}(v_1) = (\theta v_0, \mu)$$

Nous allons déterminer les constantes θ et μ . On a $[\theta v_0]^2 = \mu$.

Il s'agit maintenant de minimiser la distance :

$$d^2 = (\theta - 1)^2 \|v_0\|^2 + (\mu - \lambda_0)^2 ; \text{ ce qui donne}$$

$$d^2 = (\theta - 1)^2 \|v_0\|^2 + (\theta^2 [v_0]^2 - \lambda_0)^2. \text{ On doit avoir}$$

$$d' = 0.$$

ce qui donne :

$$\theta^3 + \theta \left(\frac{\|v_0\|^2 - 2\lambda_0 [v_0]^2}{2[v_0]^4} \right) - \frac{\|v_0\|^2}{2[v_0]^4} = 0$$

La racine qui convient est celle pour laquelle on a

$$\theta < 1 \text{ et } \theta^2 [v_0]^2 > \lambda_0.$$

I - 3. METHODES ITERATIVES REGULARISANTES.

Dans ce paragraphe, nous étudions plus particulièrement l'approximation de la solution d'équations aux dérivées partielles non linéaires. Tous ces résultats se généralisent facilement au cas des inégalités variationnelles non linéaires.

On se propose de résoudre dans un espace de dimension finie V_h l'équation

$$(3.1) \quad A_h u_h = f_h \text{ pour } f_h \text{ donné dans } V'_h$$

où A_h est un opérateur non nécessairement linéaire qui applique V_h dans V'_h .

Les hypothèses faites sur l'opérateur A_h seront plus générales que celles considérées dans le théorème 1.1 de ce chapitre et dans le théorème 7.1 de [3].

Si par ailleurs on se donne un espace de Hilbert V et un opérateur A non nécessairement linéaire de V dans son dual V' ; notre méthode permettra de résoudre directement et sans passer par l'étape de discrétisation, l'équation

$$(3.2) \quad Au = f \text{ pour } f \text{ donné dans } V'$$

Dans ce cas les conditions supposées sur l'opérateur A seront les mêmes que celles que nous allons supposer sur l'opérateur A_h de l'équation (3.1).

3.1 Méthode I : Régularisation et passage à la limite.

On fait les hypothèses suivantes :

H5) Soit $S_h \in \mathcal{L}(V_h, V'_h)$ tel que

$$i) \quad (S_h u_h, v_h)_h = (u_h, S_h v_h)_h \quad \forall u_h, v_h \in V_h$$

ii) Il existe une constante $\sigma > 0$ telle que

$$(S_h u_h, u_h)_h = [u_h]^2 \geq \sigma \|u_h\|_h^2$$

$[u_h]$ sera alors une norme équivalente à $\|u_h\|_h$

H6) A_h est hemicontinu de V_h dans V'_h et vérifie :

$$(A_h u_h - A_h v_h, u_h - v_h)_h \geq (\varphi(\|u_h\|_h) - \varphi(\|v_h\|_h)) (\|u_h\|_h - \|v_h\|_h)$$

avec $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{R}$ strictement croissante et tel que

$$\lim_{r \rightarrow \infty} \varphi(r) = +\infty$$

H7) Soit $B_h : V_h \longrightarrow V'_h$ linéaire (continue bornée si $V_h = \mathcal{H}$ de dimension infinie) tel que

$$1) (B_h u_h, u_h)_h \geq k[u_h]^2, \quad k > 0$$

ii) quelle que soit la constante $N > 0 \exists C(N)$ tel que si $[u_h] \leq N, [v_h] \leq N$, on a pour tout $\lambda > 0$:

$$(A_h u_h - A_h v_h, w_h)_h + \lambda (B_h(u_h - v_h), w_h) \leq C(N) [u_h - v_h] [w_h] \quad \forall w_h \in V_h$$

Nous avons alors le

Théorème 3.1

On suppose H5), H6), H7) alors

1°) L'équation (3.1) admet une solution unique $u_h \in V_h$.

2°) La suite $(u_{h,p}^n)$ définie par l'itération

$$(3.3) \quad u_{h,p}^{n+1} = u_{h,p}^n - \rho S_h^{-1} (A_h u_{h,p}^n + \varepsilon_p B_h u_{h,p}^n - f_h - \varepsilon_p g_h)$$

converge dans V_h vers u_h quand $n \longrightarrow \infty$ et $p \longrightarrow \infty$ pour tout g_h donné dans V'_h (ε_p étant une suite de réels tendant vers 0 quand $p \longrightarrow \infty$), pour $u_h^0 = 0$ et $\rho = k / C^2(N)$.

Démonstration

1°) L'hypothèse H6) assure l'existence et l'unicité de la solution $u_h \in V_h$ de (3.1).

2°) Fixons l'entier p . Alors l'opérateur $A_h + \varepsilon_p B_h$ vérifie les hypothèses du théorème 1.1 (il suffit de faire $X = V_h$). Donc la suite $(u_{h,p}^n)$ définie par (3.3) converge vers la solution $u_{h,p}$ de l'équation

$$(3.4) \quad A_h u_{h,p} + \varepsilon_p B_h u_{h,p} = f_h + \varepsilon_p g_h$$

et d'après le théorème 4.1 de [14] on a :

$$u_{h,p} \longrightarrow u_h \quad \text{quand } p \longrightarrow \infty, \quad u_h \text{ étant la solution de l'équation (3.1).}$$

Remarque 3.1

1°) Le théorème 3.1 est encore valable si on remplace V_h par un espace de Hilbert et l'étape de discrétisation est inutile.

Nous donnons dans le chap. II paragraphe II-1 un exemple pour illustrer

cette situation. Nous verrons que bien que l'on soit obligé d'opérer une certaine "discrétisation" au stade de l'exploitation numérique, on peut prendre le paramètre h aussi petit que l'on veut (suivant les possibilités de la machine) et l'erreur de discrétisation devient ainsi nulle. En général on obtient une très haute précision.

2°) Les itérations (3.3) sont indépendantes des conditions aux limites (Dirichlet, Neumann ou mixtes) pourvu que l'on ait un schéma de discrétisation convergent, ce qui rend inutile de recourir aux opérateurs p_h et r_h d'AUBIN (cf[1]) qui sont très difficiles à expliciter dans le cas de conditions aux limites de types Neumann par exemple.

Remarque 3.2

L'algorithme de calcul (3.3) consiste à fixer p et à itérer en n jusqu'à la solution $u_{h,p}$ de (3.4), en prenant $u_{0,p} = 0$ par exemple,

$$N = \frac{2}{k\sqrt{\epsilon}} \|f_h + \epsilon_p g_h - A_h(0) - \epsilon_p B_h(0)\|_h \text{ et } \rho = k/C^2(N).$$

Puis après avoir diminué ϵ_p on reprend les itérations (3.3). Bien entendu on peut faire le tout en une seule itération en prenant $\epsilon_p = \epsilon_n \longrightarrow 0$ quand $n \longrightarrow \infty$ et la formule

$$(3.5) \quad u_h^{n+1} = u_h^n - \rho S_h^{-1}(A_h u_h^n + \epsilon_n B_h u_h^n - f_h - \epsilon_n g_h)$$

mais des considérations techniques sur la rapidité de convergence conduisent à préférer la méthode (3.3) avec $\epsilon_p \neq \epsilon_n$. En effet nous avons :

$$(3.6) \quad [u_{h,p}^{n+1} - u_{h,p}^n]^2 \leq \theta_p [u_{h,p}^n - u_{h,p}^{n-1}]^2 \text{ avec } \theta_p = 1 - \epsilon_p^2 / C^2(N).$$

Pour que θ_p ne soit pas dès le début des itérations, très voisin de l'unité ; il est nécessaire dans le calcul des premières solutions $u_{h,p}$ ($p = 0, 1, 2 \dots$) que le paramètre ϵ_p ne soit pas trop petit.

Remarque 3.3

1°) Si nous prenons à priori $g_h = B_h u_h$ les équations (3.1) et (3.4) deviennent identiques et en posant $\epsilon_p = \epsilon_n$, les itérations

$$(3.7) \quad u_h^{n+1} = u_h^n - \rho S_h^{-1}(A_h u_h^n + \epsilon_n B_h u_h^n - f_h - \epsilon_n B_h u_h^n)$$

conduisent directement à la solution u_h de (3.1).

Ceci montre l'intérêt technique de l'estimation à priori de $B_h u_h$.

L'opérateur B_h choisi devra donc vérifier H7) et permettre si possible cette estimation.

2°) On aura aussi intérêt à utiliser $\rho = \rho_n$ cf. les remarques de I-1, 1.2.

Remarque 3.4

Toujours dans le cadre du théorème 3.1, on peut relier l'erreur sur la discrétisation $\|p_h u_h - u\| = \delta(h)$ à l'erreur sur l'itération. Plus précisément on a :

$$\|p_h u_h^n - u\| \leq \|p_h u_h^n - p_h u_h\| + \|p_h u_h - u\| \leq \frac{1}{\sqrt{\theta}} [u_h^n - u_h] + \delta(h), \quad u_h^0 = 0.$$

Donc

$$(3.8) \quad \|p_h u_h^n - u\| \leq \frac{1}{\sqrt{\theta}} \theta^{\frac{n}{2}} [u_h] + \delta(h)$$

Il suffira donc de prendre

$$(3.9) \quad n \sim \frac{2 \operatorname{Log} \frac{\sqrt{\theta} \cdot \delta(h)}{[u_h]}}{\operatorname{Log} \theta} \quad \text{avec } \theta = 1 - k^2 / c^2(N)$$

pour avoir $\|p_h u_h^n - u\| \sim 0$.

3.2 Méthode II : Méthode itérative régularisante.

Nous allons donner maintenant une autre méthode itérative directement suggérée par la remarque 3.3.

Son but essentiel est de se passer de l'hypothèse (assez restrictive pour certains opérateurs non linéaires) du théorème 1.1, sur l'opérateur A_h :

$$(3.10) \quad (A_h u_h - A_h v_h, u_h - v_h)_h \geq \alpha \|u_h - v_h\|_h^2, \quad \alpha > 0$$

Plus précisément nous supposons seulement que l'opérateur A_h vérifie la condition :

$$(3.11) \quad (A_h u_h - A_h v_h, u_h - v_h)_h \geq (\varphi(\|u_h\|_h) - \varphi(\|v_h\|_h)) \cdot (\|u_h\|_h - \|v_h\|_h)$$

Cette condition est plus faible que (3.10) et entraîne l'existence et l'unicité de la solution $u_h \in V_h$ de l'équation :

$$(3.12) \quad A_h u_h = f_h \text{ pour } f_h \text{ donné dans } V'_h.$$

Dans la méthode I on avait des difficultés techniques quand $\varepsilon \rightarrow 0$. Dans la méthode II on régularise l'opérateur A_h en lui ajoutant un opérateur λB_h où λ est une constante quelconque ; si bien que la difficulté du $\varepsilon \rightarrow 0$ disparaît. Nous avons alors le

Théorème 3.2

On fait les hypothèses H5), H6) e

H8) $B_h : V_h \rightarrow V'_h$ est linéaire et tel que

$$1) \quad (B_h u_h, v_h)_h \leq M[u_h][v_h] ;$$

$$(B_h u_h, u_h)_h \geq k[u_h]^2 \text{ avec } \frac{M}{k} \leq 1$$

11) Quelle que soit la constante $N > 0$, $\exists C(N)$ tel que si

$$[u_h] \leq N, [v_h] \leq N \text{ on a pour } \lambda > 0$$

$$(A_h u_h - A_h v_h, w_h)_h + \lambda (B_h(u_h - v_h), w_h)_h \leq C(N) [u_h - v_h][w_h] \quad \forall w_h \in V_h. \text{ Alors}$$

1°) L'équation $A_h u_h = f_h$ admet une solution unique $u_h \in V_h$

2°) La suite (u_h^n) définie par l'itération

$$(3.13) \quad A_h u_h^n + \lambda B_h u_h^{n-1} = f_h + \lambda B_h u_h^{n-1}$$

converge dans V_h vers u_h .

Démonstration.

1°) L'existence et l'unicité de la solution résulte de l'hypothèse H6).

2°) L'itération (3.13) a un sens puisque l'opérateur $A_h + \lambda B_h$ vérifie les hypothèses du théorème 1.1.

Pour u_h^0 donné on peut calculer u_h^1 à l'aide de la formule :

$$(3.14) \quad u_{h,m+1}^1 = u_{h,m}^1 - \rho S_h^{-1} (A_h u_{h,m}^1 + \lambda B_h u_{h,m}^1 - f_h - \lambda B_h u_h^0)$$

en prenant $\rho = k / C^2(N)$ où k et $C(N)$ sont donnés par l'hypothèse H8) et

$$N = \frac{2}{k\sqrt{\sigma}} \|f_h + \lambda B_h u_h^0 - A_h(o) - \lambda B_h(o)\| \text{ (si } u_h^0 = 0).$$

Par ailleurs nous avons successivement :

$$A_h u_h^n + \lambda B_h u_h^n = f_h + \lambda B_h u_h^{n-1}$$

$$A_h u_h + \lambda B_h u_h = f_h + \lambda B_h u_h.$$

En posant $\varepsilon_h^n = u_h^n - u_h$ il vient :

$$(A_h u_h^n - A_h u_h, \varepsilon_h^n)_h + \lambda (B_h \varepsilon_h^n, \varepsilon_h^n)_h = \lambda (B_h \varepsilon_h^{n-1}, \varepsilon_h^n)_h$$

D'après H8) on a :

$$[\varepsilon_h^n] \leq \frac{M}{k} [\varepsilon_h^{n-1}]. \quad (1)$$

Si $\frac{M}{k} < 1$, le théorème est démontré. Si $\frac{M}{k} = 1$ alors $[\varepsilon_h^n] \leq \text{cte}$, donc $[u_h^n] \leq \text{cte} \quad \forall n$.

Donc suivant un ultrafiltre $\mathcal{U}$ on a : $u_h^n \longrightarrow \xi_h$ dans V_h . De la relation

$$(3.15) \quad (A_h u_h^n, u_h^n) = (f_h, u_h^n) + \lambda (B_h u_h^{n-1} - B_h u_h^n, u_h^n)_h \text{ avec } \lambda_n \longrightarrow 0 \text{ quand } n \longrightarrow \infty$$

on en déduit que $\lim_{\mathcal{U}} \sup (A_h u_h^n, u_h^n)_h \leq (f_h, \xi_h)_h$

Il en résulte que $A_h \xi_h = f_h$.

D'où $\xi_h = u_h$ d'après l'unicité de la solution

ce qui démontre le théorème.

(1) Si au lieu de H5) on suppose (3.10) on obtient alors l'inégalité

$$[\varepsilon_h^n] \leq \frac{M}{k+\alpha} [\varepsilon_h^{n-1}] \text{ ce qui termine la démonstration si } \frac{M}{k} \leq 1.$$

Remarque 3.5

Le théorème 3.2 est encore valable si on remplace V_h par un Hilbert $\mathcal{H}$; les opérateurs A_h et B_h étant remplacés par $A : \mathcal{H} \longrightarrow \mathcal{H}'$ et $B : \mathcal{H} \longrightarrow \mathcal{H}'$ vérifiant H5), H6 et H8). On démontre comme précédemment que l'on a :

$u^n \longrightarrow u$ dans $\mathcal{H}$ faible.

A partir de H6) on montre que $\|u^n\| \longrightarrow \|u\|$; ce qui donne $u^n \longrightarrow u$ dans $\mathcal{H}$ fort.

Dans V_h de dimension finie on a évidemment directement la convergence forte.

3.3 Méthode III : "Méthode mixte".

La méthode qui suit a les mêmes avantages que la précédente. La régularisation de l'opérateur A_h consiste essentiellement à lui ajouter l'opérateur $\frac{\epsilon}{k} I$; les paramètres ϵ et k jouant un rôle particulier. Les hypothèses sur A_h seront bien entendu moins restrictives que celles du théorème 1.1. Plus précisément nous avons le

Théorème 3.3

Soit T un paramètre fixé et N un entier. On se donne deux paramètres $\epsilon \longrightarrow 0$ et $k = \frac{T}{N} \longrightarrow 0$ de façon que $\frac{k}{\epsilon} \longrightarrow 0$. On suppose H6). De plus A_h est lipschitzien sur les ensembles bornés. Alors

1°) la suite u_h^n définie par les itérations

$$(3.16) \quad A_h u_h^n + \frac{\epsilon}{k} u_h^n = f_h + \frac{\epsilon}{k} u_h^{n-1} \quad (\text{avec } u_h^0 = 0)$$

est telle que $\frac{1}{N} \sum_{n=1}^N u_h^n \longrightarrow u_h$ dans V_h fort où u_h est la solution de l'équation

$$(3.17) \quad A_h u_h = f_h$$

2°) Pour n fixé, la suite $(u_{h,m}^n)$ définie par les itérations :

$$(3.18) \quad u_{h,m+1}^n = u_{h,m}^n - \rho \left(\left(A_h + \frac{\epsilon}{k} \right) u_{h,m}^n - f_h - \frac{\epsilon}{k} u_h^{n-1} \right)$$

converge fortement vers u_h^n quand $m \longrightarrow \infty$ pour $u_{h,0}^m$ convenable et

$$\rho = \frac{\epsilon C_1^2(h)}{k(C(N) + \frac{\epsilon}{k} C_2^2(h))^2}$$

où $C(N)$ est la constante de Lipschitz de l'opérateur A_h (on peut prendre, si $u_{h,0}^m = 0$;

$N = \frac{2k}{\epsilon} \|f_h + \frac{\epsilon}{k} u_h^{n-1} - A_h(o)\|$) et $C_1(h)$ et $C_2(h)$ sont deux constantes assurant

l'équivalence des normes : $C_1(h) \|u_h\|_h \leq |u_h|_h \leq C_2(h) \|u_h\|_h$.

Démonstration

1°) On se donne deux espaces de Hilbert V et H $V \subset H$ avec injection continue on muni V de la norme $\|\cdot\|$ et H de la norme $|\cdot|$. Nous allons démontrer le théorème dans le cas où $V_h = V =$ espace de Banach.

a) l'équation (3.16) s'écrit alors

$$(3.19) \quad Au^n + \frac{\varepsilon}{k} u^n = f + \frac{\varepsilon}{k} u^{n-1} \quad (u^0 = 0) \text{ avec } f \text{ donné dans } V'.$$

Or (3.19) équivaut à

$$(3.20) \quad (Au^n, v - u^n) + \frac{\varepsilon}{k} (u^n - u^{n-1}, v - u^n) \geq (f, v - u^n) \quad \forall v \in V$$

Faisons $v = 0$ dans (3.20), on a :

$$- (Au^n, u^n) - \frac{\varepsilon}{2k} (|u^n|^2 - |u^{n-1}|^2 + |u^n - u^{n-1}|^2) \geq - (f, u^n)$$

or

$$(Au - A(0), u) \geq (\varphi(\|u\|) - \varphi(0)) \|u\|$$

Donc

$$\frac{\varepsilon}{2k} (|u^n|^2 - |u^{n-1}|^2 + |u^n - u^{n-1}|^2) \leq \|u^n\| (\|f\|^* + \varphi(0) + \|A(0)\|^*) - \|u^n\| \varphi(\|u^n\|)$$

or l'application $r \rightarrow \alpha r - r \varphi(r)$ est bornée dans R_+ pour $\alpha \in R$. D'où

$$(3.21) \quad \frac{\varepsilon}{2k} (|u^n|^2 - |u^{n-1}|^2 + |u^n - u^{n-1}|^2) \leq C_1.$$

On en déduit

$$(3.22) \quad \frac{\varepsilon}{2k} |u^n|^2 \leq C_2 = C_1 N \quad \text{et}$$

$$(3.23) \quad \frac{\varepsilon}{2k} \sum_{n=1}^N |u^n - u^{n-1}|^2 \leq C_2 = C_1 N$$

b) posons $u_\varepsilon = \frac{1}{N} \sum_{n=1}^N u^n$ et montrons que $\|u_\varepsilon\|$ est borné.

La relation (3.20) équivaut à

$$(3.24) \quad (A v, v - u^n) + \frac{\varepsilon}{k} (u^n - u^{n-1}, v - u^n) \geq (f, v - u^n) \quad \forall v \in V$$

Sommons (3.24) de $n=1$ à $n=N$ divisons par N et faisons $v = \frac{u_\varepsilon}{2}$, on a :

$$(A \frac{u_\varepsilon}{2}, \frac{u_\varepsilon}{2}) \leq (f, \frac{u_\varepsilon}{2}) - \frac{\varepsilon}{2T} \sum_{n=1}^N (|u^n|^2 - |u^{n-1}|^2 + |u^n - u^{n-1}|^2) + \frac{\varepsilon}{2T} \sum_{n=1}^N (u^n - u^{n-1}, u_\varepsilon)$$

ce qui donne :

$$\frac{(A \frac{u_\epsilon}{2}, \frac{u_\epsilon}{2})}{\|\frac{u_\epsilon}{2}\|} \leq C_3$$

D'où $\|u_\epsilon\| \leq C_4$

c) montrons que $u_\epsilon \longrightarrow u$ dans V fort u étant la solution de $Au = f$

Suivant un ultrafiltre $\mathcal{U}$ on a $u_\epsilon \longrightarrow \xi$ dans V faible quand $\epsilon, k \longrightarrow 0$ avec $k/\epsilon \longrightarrow 0$.

En passant à la limite dans

$$(3.25) \quad (Av, v - u_\epsilon) + \frac{\epsilon}{T} \sum_{n=1}^N (u^n - u^{n-1}, v) - \frac{\epsilon}{2T} \sum_n (|u^n|^2 - |u^{n-1}|^2 + |u^n - u^{n-1}|^2) \\ \geq (f, v - u_\epsilon)$$

On a compte tenu de (3.21), (3.22), (3.23) :

$$(Av, v - \xi) \geq (f, v - \xi) \quad \forall v \in V$$

ce qui équivaut à

$$(3.26) \quad A\xi = f$$

comme l'équation $Au = f$ admet une solution unique $u \in V$; il résulte que $\xi = u$.

Puisque $u_\epsilon \longrightarrow u$ dans V faible alors $w_\epsilon = \frac{u_\epsilon + u}{2} \longrightarrow u$ dans V faible.

Posons $v = w_\epsilon$ dans (3.25); on en déduit $\limsup (Aw_\epsilon, w_\epsilon - u) \leq 0$.

Or

$$(\varphi(\|w_\epsilon\|) - \varphi(\|u\|)) (\|w_\epsilon\| - \|u\|) \leq (Aw_\epsilon - Au, w_\epsilon - u) \\ \leq (Aw_\epsilon, w_\epsilon - u) - (Au, w_\epsilon - u) \longrightarrow 0$$

Donc

$$w_\epsilon \longrightarrow u \quad \text{et} \quad \|w_\epsilon\| \longrightarrow \|u\|$$

et alors $w_\epsilon \longrightarrow u$ dans V fort

Il en résulte que $u_\epsilon \longrightarrow u$ dans V fort

2°) On applique les itérations du théorème 1.1 à l'équation (3.16), où n est fixé.

Remarque 3.6

Dans les théorèmes 3.1 et 3.2 les constantes k et $C(N)$, N et par conséquent le paramètre ρ dépendent du pas h de discrétisation. Ceci provient d'équivalence de deux normes $\| \cdot \|_h$ et $| \cdot |_h$ sur l'espace V_h .

En général les constantes $C_i(h) \longrightarrow \infty$ quand $h \longrightarrow 0$.

Dans la pratique le problème discrétisé est résolu au stade numérique avec h petit et fixé à l'avance. Il ne reste plus donc qu'à faire le passage à la limite pour n grand.

Remarque 3.7 Si on suppose $\ell/k \longrightarrow 0$ quand $n \longrightarrow \infty$

On peut montrer que la suite (u^n) définie par (3.19) converge fortement vers u solution de l'équation $Au = f$.

Nous avons en effet

$$(3.27) \quad (Au^n, u^n) \leq (f, u^n) - \frac{\varepsilon}{2k} (|u^n|^2 - |u^{n-1}|^2 + |u^n - u^{n-1}|^2)$$

De (3.27) et (3.21) on déduit que u^n est borné.

Donc suivant un ultrafiltre $\mathcal{U}$ on a $u^n \longrightarrow \xi$ dans V faible. Or (3.20) peut encore s'écrire :

$$(3.28) \quad (Av, v - u^n) + \frac{\varepsilon}{k} (u^n - u^{n-1}, v) - \frac{\varepsilon}{2k} (|u^n|^2 - |u^{n-1}|^2 + |u^n - u^{n-1}|^2) \\ \geq (f, v - u^n)$$

En passant à la limite suivant $\mathcal{U}$ on a

$$(Av, v - \xi) \geq (f, v - \xi) \quad \forall v \in X \quad \text{D'où}$$

$$(3.29) \quad A\xi = f \quad \text{et} \quad \xi = u$$

On montre encore que

$$\limsup (Au^n, u^n) \leq (f, u).$$

ce qui donne :

$$(\varphi(\|u^n\|) - \varphi(\|u\|)) (\|u^n\| - \|u\|) \leq (Au^n - Au, u^n - u) \longrightarrow 0.$$

Donc

$$\|u^n\| \longrightarrow \|u\|$$

D'où

$$u^n \longrightarrow u \text{ dans } V \text{ fort.}$$

II - EQUATIONS AUX DERIVEES PARTIELLES NON LINEAIRES ET PROBLEMES

D'OPTIMISATION SUR L'ESPACE ENTIER.

II - 1 Problème 1

Soit $H_0^1(0,1) = \{u | u \in L^2(0,1), \frac{du}{dx} \in L^2(0,1); u(0) = u(1) = 0\}$

la dérivée étant prise au sens des distributions. On désigne par $H^{-1}(0,1)$ le dual de $H_0^1(0,1)$ et on munit $H_0^1(0,1)$ de la norme

$$\|u\| = \left\| \frac{du}{dx} \right\|_{L^2}.$$

On cherche $u \in H_0^1(0,1)$ vérifiant

$$(2.1) \quad u'' = f(x, u) \quad (1)$$

on fait les hypothèses suivantes

$$i) \quad f(x, u) \in \mathcal{B}^0([0,1] \times \mathbb{R})$$

$$ii) \quad \text{Il existe une constante } \gamma > 0 \text{ telle que}$$

$$\frac{f(x, u_1) - f(x, u_2)}{u_1 - u_2} \geq -\gamma > -\pi^2; \quad u_1 \neq u_2, \quad \forall x \in [0,1] \text{ et } \gamma \geq 0.$$

$$iii) \quad \text{Quelque soit } N > 0, \exists \text{ une constante } C(N) \text{ telle que si}$$

$$|u_1| \leq N, \quad |u_2| \leq N \text{ on a}$$

$$\frac{f(x, u_1) - f(x, u_2)}{u_1 - u_2} \leq \alpha(N) < +\infty \text{ pour } u_1 \neq u_2, \quad \forall x \in [0,1].$$

(1) Ce problème a été étudié dans [7] par d'autres méthodes.

Nous avons alors le

Théorème 2.1

On suppose i), ii) et iii) alors

1 - Il existe $u \in H_0^1(0,1)$ unique tel que $u'' = f(x, u)$

2 - Posons $A u = -u'' + f(x, u)$; $S u = -\frac{d^2 u}{dx^2}$ alors la suite (u_n) définie par les itérations :

$$(2.2) \quad u_{n+1} = u_n - \rho S^{-1} A u_n$$

converge dans $H_0^1(0,1)$ fort vers la solution u de (2.1) pour $u_0 = 0$ et $\rho > 0$ convenable donné par (2.5).

Démonstration

L'opérateur $A u = -u'' + f(x, u)$ applique $H_0^1(0,1)$ dans $H^{-1}(0,1)$.

Vérifions les hypothèses du théorème 1.1 paragraphe I-1.

Posons $[u] = \|u\|$. D'après (i) on a :

$$1 - (A u - A v, u - v) \geq \|u' - v'\|_{L^2(0,1)}^2 - \gamma \|u - v\|_{L^2(0,1)}^2$$

Il résulte du calcul des variations que

$$\pi^2 = \inf_{v \in H_0^1(0,1)} \frac{\int_0^1 |v'|^2 dx}{\int_0^1 |v|^2 dx}$$

Donc $\|v\|_{L^2(0,1)}^2 \leq \frac{1}{\pi^2} \|v'\|_{L^2(0,1)}^2 \quad \forall v \in H_0^1(0,1)$ D'où

$$(2.3) \quad (A u - A v, u - v) \geq k [u - v]^2 \quad \text{avec} \quad k = 1 - \frac{\gamma}{\pi^2} ; 0 \leq \gamma < \pi^2$$

2 - D'après ii) et iii), si $|u| \leq N$, $|v| \leq N$ nous avons

$$|f(x, u) - f(x, v)| \leq |u - v| \sup (|\alpha(N)|, |\gamma|).$$

Supposons que $[u] \leq N$ et $[v] \leq N$. Alors

$$|(Au - Av, w)| \leq \|u' - v'\|_{L^2}^2 \|w'\|_{L^2}^2 + \|f(x, u) - f(x, v)\| \|w\|.$$

D'où

$$(2.4) \quad |(Au - Av, w)| \leq C(N) [u - v][w] \quad \text{avec} \quad C(N) = 1 + \frac{1}{\pi^2} \sup(|\alpha|, |\gamma|)$$

Les hypothèses du théorème 1.1 chap. I sont donc vérifiées. Si on pose :

$$X = H_0^1(0,1) ; \quad u_0 = 0, \quad N = \frac{2}{k} \|A(0)\|_{L^2}^2 \quad \text{et} \quad \rho = k/C^2(N)$$

Remarque 2.1

Les itérations (2.2) convergent pourvu que ρ vérifie :

$$(2.5) \quad 0 < \rho < \frac{2k}{C^2(N)} \quad \text{où} \quad k = 1 - \frac{\gamma}{\pi^2} ; \quad C(N) = 1 + \frac{1}{\pi^2} \sup(|\alpha|, |\gamma|) ; \quad 0 \leq \gamma < \pi^2$$

Nous allons montrer maintenant que la discrétisation n'introduit aucune erreur dans le problème et pourvu que l'on prenne un paramètre $h > 0$ suffisamment petit ; tout se passe comme si on itérait effectivement dans l'espace de l'Hilbert $H_0^1(0,1)$. Ceci est d'ailleurs valable chaque fois que le problème aux limites est bien posé dans un Hilbert quelconque $\mathcal{H}$.

Soit N un entier et h un paramètre > 0 tel que $Nh = 1$. On divise l'intervalle $(0,1)$ en N intervalles $(0,h), (h,2h), \dots, (N-1)h, Nh$

Soit θ_h^i la fonction caractéristique de $]ih, (i+1)h[$; $i = 0, \dots, N-1$.

Posons :

$$V_h = \{u_h \mid u_h = \sum_{i=0}^{N-1} u_h^i \theta_h^i, \quad u_h^i \in \mathbb{R} ; \quad u_h^0 = u_h^{N-1} = 0\}$$

V_h est un espace de dimension finie que nous mettons en dualité avec V'_h à l'aide du produit scalaire

$$(u_h, v_h)_h = h \sum_{i=0}^{N-1} u_h^i v_h^i$$

Posons
$$\nabla_h u_h(x) = \frac{u_h(x + \frac{h}{2}) - u_h(x - \frac{h}{2})}{h} \quad ; \quad \nabla_h^2 u_h = \nabla_h(\nabla_h u_h)$$

Soit $\chi_h = \frac{1}{h} \theta_h^0$. On pose
$$p_h u_h = \sum_{i=0}^{N-1} u_h^i \chi_h \cdot \theta_h^i$$

(où le signe $\cdot$ désigne le produit de convolution)

$r_h u(x) = u(th)$, pour $x \in]th, (t+1)h[$, $t = 1, \dots, N-1$.

Proposition 2.1

Posons $f_h(x, u_h) = \hat{r}_h f(x, u)$ $u_h \in V_h$, $u \in H_0^1(0,1)$

(2.6) $A_h u_h = -\nabla_h^2 u_h + f_h(x, u_h)$ Alors

Il existe $u_h \in V_h$ unique tel que

(2.7) $A_h u_h = 0$

De plus on a $p_h u_h \longrightarrow u$ dans $H_0^1(0,1)$ fort ; u étant la solution de (2.1).

Proposition 2.2

La suite (u_h^n) définie par les itérations

(2.8) $u_h^{n+1} = u_h^n - \rho S_h^{-1} A_h u_h^n$

où $S_h = -\nabla_h^2$ et A_h donné par (2.6) converge vers u_h solution de (2.7) pour $u_h^0 = 0$ et $\rho > 0$ donné par (2.9).

Démonstration

Nous avons en effet :

$$(A_h u_h - A_h v_h, u_h - v_h)_h = (A p_h u_h - A p_h v_h, p_h u_h - p_h v_h) = X$$

$$X \geq \left(1 - \frac{\gamma}{\pi^2}\right) \left\| \frac{d}{dx} (p_h u_h - p_h v_h) \right\|_{L^2}^2 = \left(1 - \frac{\gamma}{\pi^2}\right) \left\| \nabla_h u_h - \nabla_h v_h \right\|_{L^2}^2$$

D'où

$$(A_h u_h - A_h v_h, u_h - v_h)_h \geq k [u_h - v_h]^2 \text{ avec } k = 1 - \frac{\gamma}{\pi^2}$$

Si on pose $[u_h] = \left\| \nabla_h u_h \right\|_{L^2}$.

De même on a :

$$(A_h u_h - A_h v_h, w_h) \leq C(N) [u_h - v_h] [w_h]$$

$$\text{où } C(N) = 1 + \frac{1}{2} \sup (|\alpha|, |\gamma|).$$

Il suffit donc de poser

$$(2.9) \quad u_h^0 = 0, \quad N = \frac{2}{k} \|A_h(0)\|_{L^2}^2 \quad \text{et} \quad \rho = k/C^2(N).$$

Remarque 2.2

1 - Nous avons donc $u_h^n \longrightarrow u_h$ dans V_h et

$p_h u_h \longrightarrow u$ dans $H_0^1(0,1)$ fort ; u étant la solution de (2.1).

2 - Comme le paramètre ρ est indépendant de h , il résulte que l'on peut prendre h aussi petit que l'on veut ; ce qui annule l'erreur de discrétisation.

3 - A priori nous avons les majorations de l'erreur suivantes :

$$\left\| u_h^n - u_h \right\|_{L^2}^2 \leq \frac{1}{C_i} \left\| \nabla_h u_h^n - \nabla_h u_h \right\|_{L^2}^2 \leq \theta^n \left\| \nabla_h u_h \right\|_{L^2}^2$$

où C_1 est une constante de l'ordre de π^2 d'après un calcul de variation "discret".

$$\theta = 1 - \frac{\kappa^2}{C^2(N)} ; \quad \kappa = 1 - \frac{\gamma}{\pi^2} ; \quad C(N) = 1 + \frac{1}{\pi^2} \sup (\alpha(N), \gamma)$$

$$\text{et} \quad N = \frac{2}{\kappa} \|f(x, 0)\|_{L^2}$$

En ce qui concerne l'erreur sur la convergence de $p_h u_h$ vers u nous avons à priori la majoration $\|p_h u_h - u\| \leq C_2 h$.

En effet d'après la remarque 5.3 chap. I de [14] nous avons

$$\|p_h u_h - u\| \leq C_3 \|p_h r_h u - u\|.$$

On montre dans [2] que si $u \in H^q(0, 1)$ on a :

$$\|p_h r_h u - u\|_{H^m(0, 1)} \leq C_u h^{q-m}, \quad q \geq m, \quad C_u > 0$$

Comme $u'' = f(x, u) \in \mathcal{C}^0([0, 1] \times \mathbb{R})$; il résulte que

$u \in H^2(0, 1)$. Donc $\|p_h u_h - u\| \leq C_2 h$.

Résultats numériques du problème 1

On cherche $u \in H_0^1(0, 1)$ vérifiant :

$$(2.10) \quad u'' = \frac{1}{2}(u + x + 1)^3 \quad x \in [0, 1]$$

Ce problème admet pour solution exacte

$$(2.11) \quad u(x) = \frac{2}{2-x} - x - 1 \quad \text{of [7]}$$

Posons

$$Au = -u'' + \frac{1}{2}(u + x + 1)^3$$

On montre facilement que

$$(Au - Av, u - v) \geq [u - v]^2 \quad \text{avec} \quad [u] = \|u'\|_{L^2(0, 1)}^2.$$

Si $[u] \leq N$ et $[v] \leq N$ on a :

$$(Au - Av, w) \leq C(N)[u - v][w] \quad \forall w \in H_0^1(0,1)$$

où $C(N) = 1 + \frac{3}{2} \left(\frac{N+2}{\pi} \right)^2$ et l'on peut prendre

$$N = 2 \left\| \frac{(x+1)^3}{2} \right\|_{L^2(0,1)}$$

Au problème (2.10) on associe alors le problème discrétisé

$$(2.12) \quad A_h u_h = 0 \quad \text{avec} \quad A_h u_h = -\nabla^2 u_h + \frac{1}{2}(u_h + u_{h+1})^3$$

$i = 0, 1, \dots, p$ où p est un entier réel tel que $ph = 1$.

Alors la suite (u_h^n) définie par l'itération

$$(2.13) \quad u_h^{n+1} = u_h^n - \rho S_h^{-1} A_h u_h^n \quad \text{avec} \quad u_h^0 = 0$$

où $S_h = -\nabla^2$ et $\rho = 1/C^2(N)$

converge vers u solution de (2.10) pour n suffisamment grand et h suffisamment petit.

Nous inversons l'opérateur S_h à l'aide de la méthode bien connue de Gauss - Seidel cf [16] en réinitialisant chaque fois à zéro, pour les itérations internes

Nous donnons ci-dessous les résultats numériques de

$$\|u_h^n - u\|_{L^2(0,1)} \quad \text{et} \quad \|u_h^n - u\|_{L^\infty(0,1)}$$

en fonction du pas h après stabilisation des résultats dans les itérations (2.13)

h	$\ u_h^n - u \ _{L^2}^2$	$\ u_h^n - u \ _{L^\infty}$
1/10	$1,965 \cdot 10^{-3}$	$2,8 \cdot 10^{-3}$
1/20	$5,214 \cdot 10^{-4}$	$6,8 \cdot 10^{-4}$
1/40	$1,325 \cdot 10^{-4}$	$1,8 \cdot 10^{-4}$
1/80	$3,356 \cdot 10^{-5}$	$4,6 \cdot 10^{-5}$

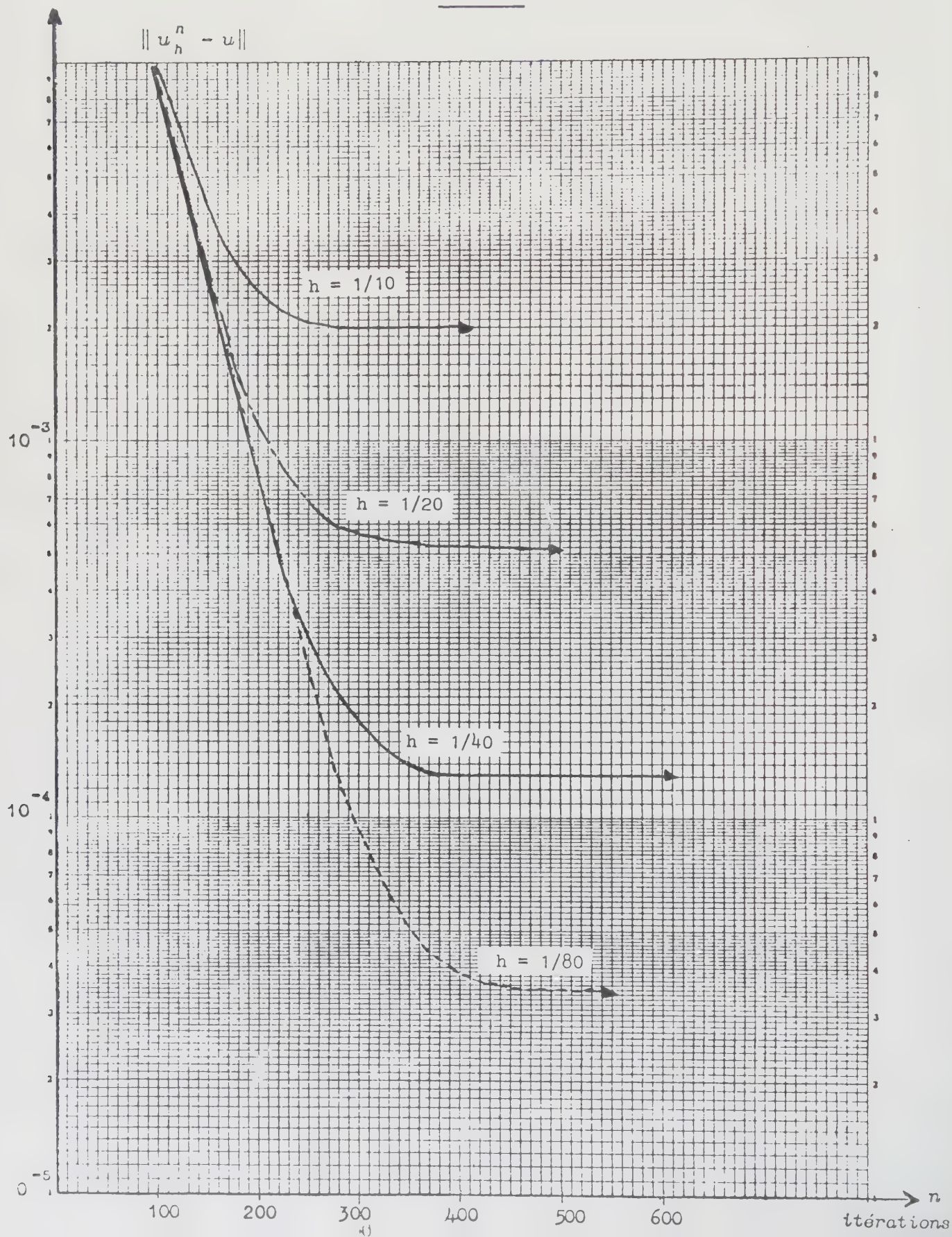
Remarque 2.3

L'exemple numérique ci-dessus a été traité par Ciarlet dans [7] par d'autres méthodes.

Notre méthode s'applique bien entendu pour des dimensions supérieures (cf problème 2).

Les courbes n° 1 donnent l'évolution de $\| u_h^n - u \|_{L^2}^2$ en fonction de n (itérations) pour $h = 1/10$, $h = 1/20$, $h = 1/40$ et $h = 1/80$.

Courbes n° 1



Temps de calcul

Le temps de calcul (compilation + exécution) pour $h = 1/20$ et 500 itérations en C.D.C. 3 600 est de 55 s. Il y a lieu de signaler que les résultats se stabilisent complètement dès la 400ème itération.

II - 2 Problème 2

Cet exemple généralise le précédent au cas de la dimension $n = 2$. La théorie est analogue à celle du problème 1.

Soit $\Omega = (0, \alpha) \times (0, \alpha)$, $\alpha > 0$ un ouvert de $\mathbb{R}^2$

On cherche $u \in H_0^1(\Omega)$ solution de

$$(2.14) \quad \Delta u = f(x, y, u(x, y)) \quad (x, y) \in \Omega$$

où f est une application non linéaire de $\Omega \times \mathbb{R}$ dans $\mathbb{R}$

on pose : $[u] = \left(\sum_{i=1}^2 \|D_i u\|_{L_2} \right)^{1/2} = (Su, u)$ avec $S = -\Delta$

pour

$$f(x, y, u) = 2 \frac{x^2 + y^2}{(\alpha - x - y)^2} \left(u + 1 + \frac{1}{\alpha - 1} (x - 1)(y - 1) \right)^3 - \frac{2(x + y)}{(\alpha - 1 - xy)^2}$$

la solution exacte de (2.14) est donnée par

$$(2.15) \quad u(x, y) = \frac{\alpha - x - y}{\alpha - 1 - xy} - 1 - \frac{1}{\alpha - 1} (x - 1)(y - 1) \quad \alpha > 2$$

Posons alors

$$(2.16) \quad Au = -\Delta u + f(x, y, u) = 0$$

au problème (2.16) on associe le problème discrétisé

$$(2.17) \quad A_h u_h = 0 \quad \text{avec} \quad h_1 = h_2 = h$$

et l'itération :

$$(2.18) \quad u_h^{n+1} = u_h^n - \rho S_h^{-1} (A u_h^n) \quad \text{avec} \quad u_h^0 = 0$$

avec

$$(2.19) \quad \rho = 1/C^2(N)^*$$

où $C(N) = 1 + \frac{\alpha(N)}{2\pi^2}$ avec $\alpha(N) = \frac{12}{(\alpha - 2)^2} \left(N + \frac{\alpha}{\alpha - 1} \right)^2$

et $N = 2 \|f(x, y, 0)\|_{L^2}$

* Le paramètre ρ est calculé comme dans le problème 1.

On sait que $u_h^n \xrightarrow{n \rightarrow \infty} u_h$ et que $p_h u_h \xrightarrow{h \rightarrow 0} u$ quand $h \rightarrow 0$ où u est la solution de (2.14).

Dans les itérations (2.18) on résoud en outre l'équation :

$$(2.20) \quad -\Delta(A_h u_h^n) = 0 = -\Delta g, \text{ en posant } g = A_h u_h^n,$$

à l'aide des itérations internes (de Gauss-Seidel cf. [16]) :

$$(2.21) \quad g_{p+1}(x, y) = \frac{1}{4} \left(g_{p+1}(x-h, y) + g_p(x+h, y) + g_{p+1}(x, y-h) + g_p(x, y+h) \right)$$

A chaque itération principale n , il est beaucoup plus rentable d'utiliser comme valeur initiale de g_0 dans (2.21) la valeur finale du calcul précédent qui peut être très voisine du nouveau résultat.

Nous prenons comme test d'arrêt pour les itérations (2.21) :

$$\frac{\|g_{p+1} - g_p\|}{\|g_{p+1}\|} < \epsilon, \quad \epsilon \text{ donné.}$$

on prendra pour $\| \cdot \| = \| \cdot \|_{L^\infty(\Omega)}$ qui nécessite moins d'opérations que $\| \cdot \|_{L_2}$ ou $\| \cdot \|_{H_0^1}$.

Influence du paramètre ϵ cf. courbes n° 2a, n° 2b et n° 2c).

Le nombre d'itérations nécessaires pour inverser le laplacien augmente bien sûr lorsqu'on diminue l'erreur tolérable ϵ . Comme le temps de calcul est directement lié au nombre d'itérations, il est avantageux de choisir ϵ pas trop petit. Afin de mettre en évidence cette valeur suffisante de ϵ , les courbes donnant $\|u_n - u\| / \|u\|$ en fonction de ϵ ont été tracées.

On constate que pour $M = \frac{1}{h} = 10$, $\alpha = 5$, $\rho = 0,796068$;

Il suffit de choisir $10^{-6} < \epsilon < 10^{-4}$. (cf tableau n° 2a)

Pour $M = \frac{1}{h} = 20$, $\alpha = 5$, $\rho = 0,985$;

il suffit de choisir $10^{-7} < \epsilon < 2 \cdot 10^{-6}$ (cf tableau n° 2b)

Un tel choix permet d'obtenir dès les premières itérations une meilleure précision que celle qui est obtenue lorsque les résultats deviennent stationnaires.

Ainsi : (cf tableau n° 2a)

pour $M = 10$ $\rho = 0,796068$, $\epsilon = 10^{-5}$

on a $\|u_n - u\| / \|u\| = 1,611 \cdot 10^{-4}$ dès l'itération 10 ;

alors que pour $\epsilon = 10^{-7}$

on a $\frac{\|u_n - u\|}{\|u\|} = 2,205 \cdot 10^{-4}$ après 40 itérations.

pour $M = 20$, $\rho = 0,985$, $\epsilon = 10^{-6}$ (tableau n° 2b)

on a : $\frac{\|u_n - u\|}{\|u\|} = 3,70 \cdot 10^{-5}$ dès la 3ème itération.

alors que pour $\epsilon = 10^{-10}$

on a $\|u_n - u\| / \|u\| = 5,84 \cdot 10^{-5}$ après 90 itérations.

Remarque 2.4

Il est possible de déterminer cette valeur suffisante de ϵ en calculant $D_\epsilon = \|u_2 - u_1\| / \|u_2\|$ pour $\epsilon = 10^{-1}$, 10^{-2} , 10^{-3} , ...

Lorsque $\frac{D_\epsilon - D_{\epsilon_1}}{D_{\epsilon_1}} < 10^{-2}$, la valeur de ϵ est suffisante.

Tableau n° 2a

Erreur relative totale $\|u_n - u\| / \|u\|$
 $M = 10$
 $\alpha = 5$
 $p = 0,796068$

ϵ	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
1	$5.673.10^{-1}$	$2.626.10^{-1}$	$2.005.10^{-1}$	$1.944.10^{-1}$	$1.937.10^{-1}$	$1.937.10^{-1}$	$1.936.10^{-1}$	$1.936.10^{-1}$	$1.936.10^{-1}$	$1.936.10^{-1}$
2	3.898	$7.023.10^{-2}$	$5.220.10^{-3}$	$6.610.10^{-3}$	$7.382.10^{-3}$	$7.452.10^{-3}$	$7.460.10^{-3}$	$7.461.10^{-3}$	$7.461.10^{-3}$	$7.461.10^{-3}$
5	3.165	5.120	3.531	$4.126.10^{-4}$	$3.954.10^{-4}$	$4.786.10^{-4}$	$4.866.10^{-4}$	$4.874.10^{-4}$	$4.875.10^{-4}$	$4.875.10^{-4}$
10	1.918	3.041	2.224	3.693	1.611	2.197	2.197	2.205	2.206	2.206
15	1.158	1.828	1.305	1.632	1.844	2.164	2.200	2.205	2.205	2.205
30	$2.540.10^{-2}$	$4.114.10^{-3}$	$1.641.10^{-4}$	1.538	2.125	2.196	2.204	2.205	2.205	2.205
45	$5.573.10^{-3}$	1.055	1.544	2.055	2.188	2.203	2.205	2.205	2.205	2.205

Erreur relative Cauchy $\|u_n - u_{n-1}\| / \|u_n\|$

ϵ	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
2	$2.200.10^{-1}$	$1.748.10^{-1}$	$1.693.10^{-1}$	$1.634.10^{-1}$	$1.633.10^{-1}$	$1.621.10^{-1}$	$1.621.10^{-1}$	$1.621.10^{-1}$	$1.621.10^{-1}$	$1.621.10^{-1}$
3	$9.122.10^{-2}$	$3.980.10^{-2}$	$3.337.10^{-2}$	$3.069.10^{-2}$	$3.044.10^{-2}$	$3.042.10^{-2}$	$3.042.10^{-2}$	$3.041.10^{-2}$	$3.041.10^{-2}$	$3.041.10^{-2}$
5	4.907	$7.152.10^{-3}$	$1.197.10^{-3}$	$1.127.10^{-3}$	$1.136.10^{-3}$	$1.133.10^{-3}$	$1.132.10^{-3}$	$1.132.10^{-3}$	$1.132.10^{-3}$	$1.132.10^{-3}$
10	2.500	3.385	$2.750.10^{-4}$	$5.859.10^{-5}$	$6.348.10^{-6}$	$3.897.10^{-7}$	$2.573.10^{-7}$	$3.110.10^{-7}$	$3.108.10^{-7}$	$3.107.10^{-7}$
15	1.390	1.987	1.563	3.635	4.159	4.673	$5.891.10^{-8}$	$6.128.10^{-9}$	$6.652.10^{-10}$	$2.446.10^{-11}$
30	$2.767.10^{-3}$	$4.210.10^{-4}$	$3.649.10^{-5}$	$8.099.10^{-6}$	$9.300.10^{-7}$	1.051	1.339	1.416	1.757	1.583
45	$5.899.10^{-4}$	$9.109.10^{-5}$	$8.041.10^{-6}$	1.765	2.027	$2.292.10^{-8}$	$2.928.10^{-9}$	$3.096.10^{-10}$	$4.419.10^{-11}$	$2.689.10^{-11}$
nombre d'itérations	>98	>115	>138	≥200	271	267	495	655	764	1107

Tableau n° 2b

Erreur relative totale : $\|u_n - u\| / \|u\|$

$M = 20$
 $\alpha = 5$
 $\rho = 0.985$

iter.	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
1	$7.9756 \cdot 10^{-1}$	$2.8651 \cdot 10^{-1}$	$4.0171 \cdot 10^{-2}$	$7.4397 \cdot 10^{-3}$	$4.6624 \cdot 10^{-3}$	$4.4446 \cdot 10^{-3}$	$4.4238 \cdot 10^{-3}$	$4.4217 \cdot 10^{-3}$	$4.4215 \cdot 10^{-3}$	$4.4215 \cdot 10^{-3}$
3	7.5620	2.6257	2.4940	3.8204	$3.3869 \cdot 10^{-4}$	$3.7099 \cdot 10^{-5}$	$5.5521 \cdot 10^{-5}$	$5.8178 \cdot 10^{-5}$	$5.8480 \cdot 10^{-5}$	$5.7510 \cdot 10^{-5}$
5	7.1951	2.4977	2.3587	3.6341	3.2037	3.7655	5.5602	5.8182	5.8424	5.8443
10	6.3562	2.2042	2.0571	3.2038	2.7694	3.8524	5.5910	5.8211	5.8427	5.8448
15	5.6160	1.9451	1.7986	2.8250	2.3874	3.9825	5.6191	5.8237	5.8429	5.8448
30	3.8720	1.3362	1.2140	1.9344	1.4976	4.4382	5.6873	5.8301	5.8435	5.8449
45	2.6662	$9.1759 \cdot 10^{-2}$	$8.2687 \cdot 10^{-3}$	1.3183	$9.0140 \cdot 10^{-5}$	4.8370	7.7358	5.8346	5.8440	5.8449
90	$8.6575 \cdot 10^{-2}$	2.9710	2.6823	$3.9602 \cdot 10^{-4}$	3.0966	5.5050		5.8416	5.8447	5.8450

Erreur relative Cauchy : $\|u_n - u_{n-1}\| / \|u_n\|$

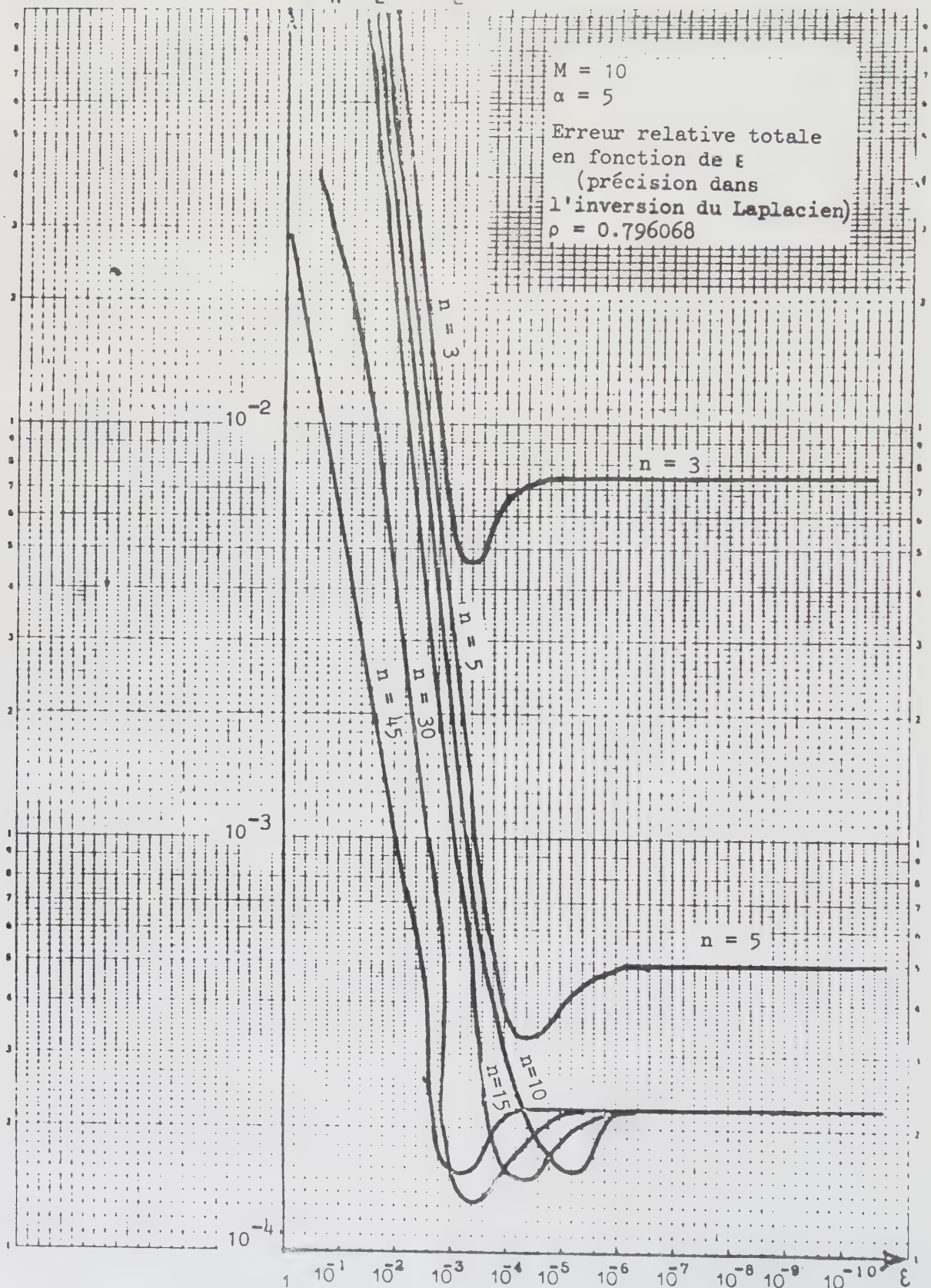
iter.	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
2	$9.5138 \cdot 10^{-2}$	$2.3539 \cdot 10^{-2}$	$1.5368 \cdot 10^{-2}$	$1.0566 \cdot 10^{-2}$	$4.8475 \cdot 10^{-3}$	$4.4077 \cdot 10^{-3}$	$4.3673 \cdot 10^{-3}$	$4.3633 \cdot 10^{-3}$	$4.3629 \cdot 10^{-3}$	$4.3629 \cdot 10^{-3}$
3	$7.5058 \cdot 10^{-2}$	$9.1339 \cdot 10^{-3}$	$9.4661 \cdot 10^{-4}$	$1.0655 \cdot 10^{-4}$	$6.0576 \cdot 10^{-5}$	$5.3580 \cdot 10^{-5}$	$5.0346 \cdot 10^{-5}$	$5.0046 \cdot 10^{-5}$	$5.0037 \cdot 10^{-5}$	$5.0036 \cdot 10^{-5}$
5	$6.2042 \cdot 10^{-2}$	$8.3567 \cdot 10^{-3}$	$7.0647 \cdot 10^{-4}$	$9.5940 \cdot 10^{-5}$	$9.3959 \cdot 10^{-6}$	$9.6777 \cdot 10^{-7}$	$9.2163 \cdot 10^{-8}$	$8.9008 \cdot 10^{-9}$	$8.4729 \cdot 10^{-9}$	$8.3465 \cdot 10^{-9}$
10	$4.2419 \cdot 10^{-2}$	$7.1102 \cdot 10^{-3}$	$5.9147 \cdot 10^{-4}$	$8.3535 \cdot 10^{-5}$	$8.2821 \cdot 10^{-6}$	$8.36 \cdot 10^{-7}$	$8.6141 \cdot 10^{-8}$	$8.7430 \cdot 10^{-9}$	$9.2981 \cdot 10^{-10}$	$1.3006 \cdot 10^{-10}$
15	$3.1334 \cdot 10^{-2}$	$6.0823 \cdot 10^{-3}$	$5.0364 \cdot 10^{-4}$	$7.3121 \cdot 10^{-5}$	$7.2983 \cdot 10^{-6}$	$7.2269 \cdot 10^{-7}$	$7.47 \cdot 10^{-8}$	$7.4715 \cdot 10^{-9}$	$7.9245 \cdot 10^{-10}$	$1.1966 \cdot 10^{-10}$
30	$1.5675 \cdot 10^{-2}$	$3.8976 \cdot 10^{-3}$	$3.2405 \cdot 10^{-4}$	$4.9957 \cdot 10^{-5}$	$5.0006 \cdot 10^{-6}$	$4.7893 \cdot 10^{-7}$	$4.9609 \cdot 10^{-8}$	$4.9433 \cdot 10^{-9}$	$5.0485 \cdot 10^{-10}$	$9.9656 \cdot 10^{-11}$
45	$5.0998 \cdot 10^{-3}$	$2.5530 \cdot 10^{-3}$	$2.1508 \cdot 10^{-4}$	$3.4465 \cdot 10^{-5}$	$3.4284 \cdot 10^{-6}$	$3.2308 \cdot 10^{-7}$	$3.3484 \cdot 10^{-8}$	$3.2317 \cdot 10^{-9}$	$3.3581 \cdot 10^{-10}$	$9.3306 \cdot 10^{-11}$
90	$2.3986 \cdot 10^{-3}$	$7.7619 \cdot 10^{-4}$	$6.7285 \cdot 10^{-5}$	$1.4232 \cdot 10^{-5}$	$1.1068 \cdot 10^{-6}$	$1.0271 \cdot 10^{-7}$		$1.0227 \cdot 10^{-9}$	$1.2895 \cdot 10^{-10}$	$7.9316 \cdot 10^{-11}$

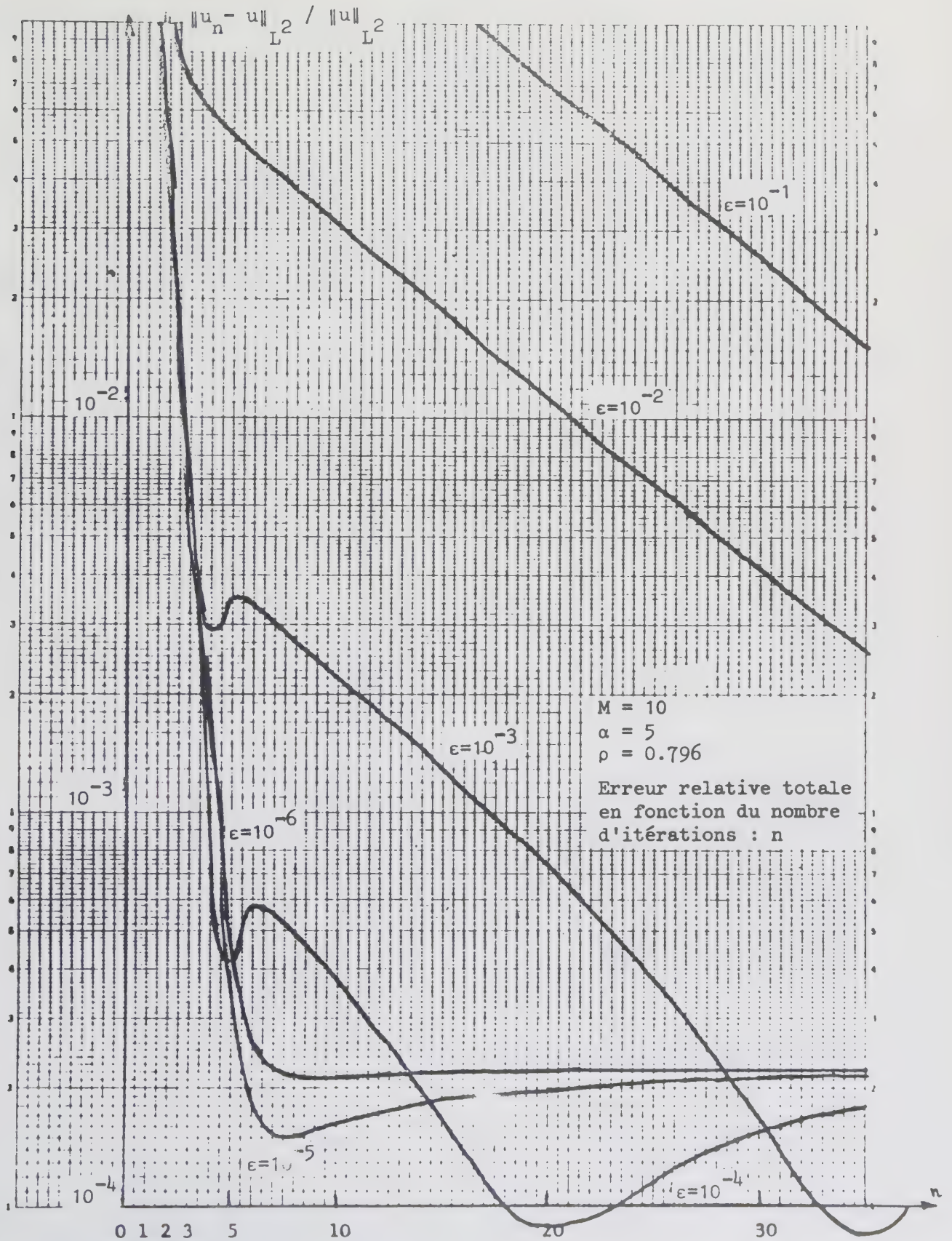
$$\|u_n - u\|_L^2 / \|u\|_L^2$$

$$M = 10$$

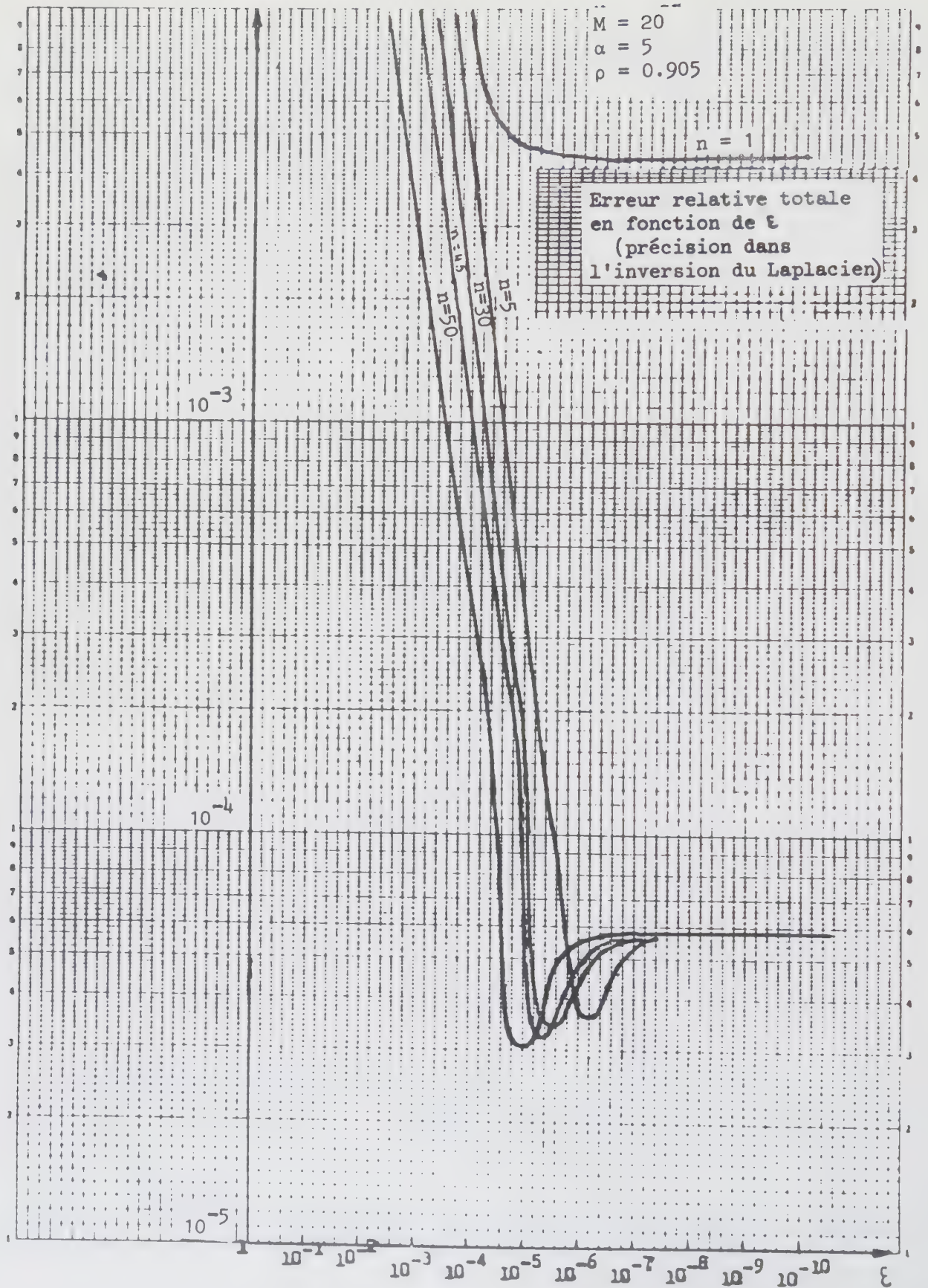
$$\alpha = 5$$

Erreur relative totale
en fonction de ε
(précision dans
l'inversion du Laplacien)
 $\rho = 0.796068$





$$\|u_n - u\|_{H_0^1} / \|u_1\|_{H_0^1}$$



Influence du paramètre ρ - (cf. courbes n° 2d, n° 2e et tableau n° 2c).

On a constaté que $\|u_n - u_{n-1}\| / \|u_n\|$ décroît d'autant plus vite que ρ est plus proche de la valeur de ρ optimale, pour décroître plus lentement dès que la stabilisation est obtenue.

Pour ρ petit la stabilisation est obtenue tard

Pour ρ très grand ($\rho \approx 2$) $\|u_n - u\| / \|u\|$ décroît puis oscille sans se stabiliser.

Entre ces cas extrême on peut obtenir une valeur de ρ optimale pour laquelle la stabilisation a lieu dès les premières itérations et sans oscillation.

Pour $M = 10$, $\alpha = 5$, $\epsilon = 10^{-8}$, $\rho = 0,987$ la stabilisation est obtenue dès la 4ème itération.

On est donc amené à chercher de façon plus précise cette valeur optimale de ρ (cf. chap. I proposition 1.1).

Partant de $u_0 = 0$

$$u_n = u_{n-1} - \rho(u_{n-1} - g_{n-1}) \quad \text{avec} \quad \Delta g_{n-1} = f(x, y, u_{n-1})$$

donne $u_1 = \rho g_0$ avec $\Delta g_0 = f(x, y, 0)$

donc $\|u_1 - u\| / \|u\| = \varphi(\rho)$ est une hyperbole

Le minimum ρ de $\|u_1 - u\| / \|u\|$ est atteint pour

$$\rho = (g_0, u) / (g_0, g_0) .$$

Ne connaissant pas u on est amené à rechercher une valeur approchée de ce minimum.

On ne peut en fait que rechercher le minimum de $\|u_2 - u_1\| / \|u_2\|$ ce qui donne une localisation assez précise de ρ_{opt} .

Tableau 2c

$M = 10$

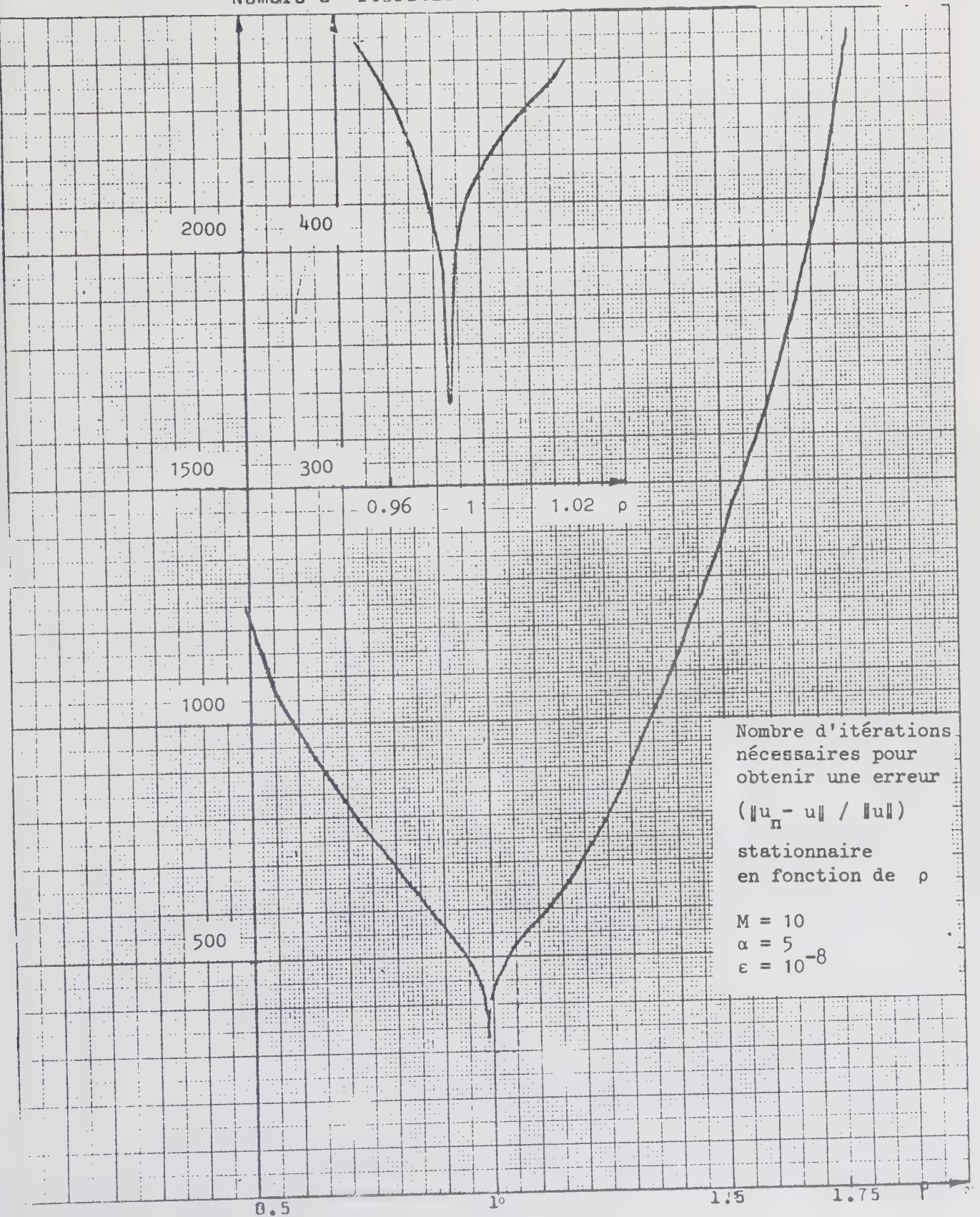
$\alpha = 5$

$\varepsilon = 10$

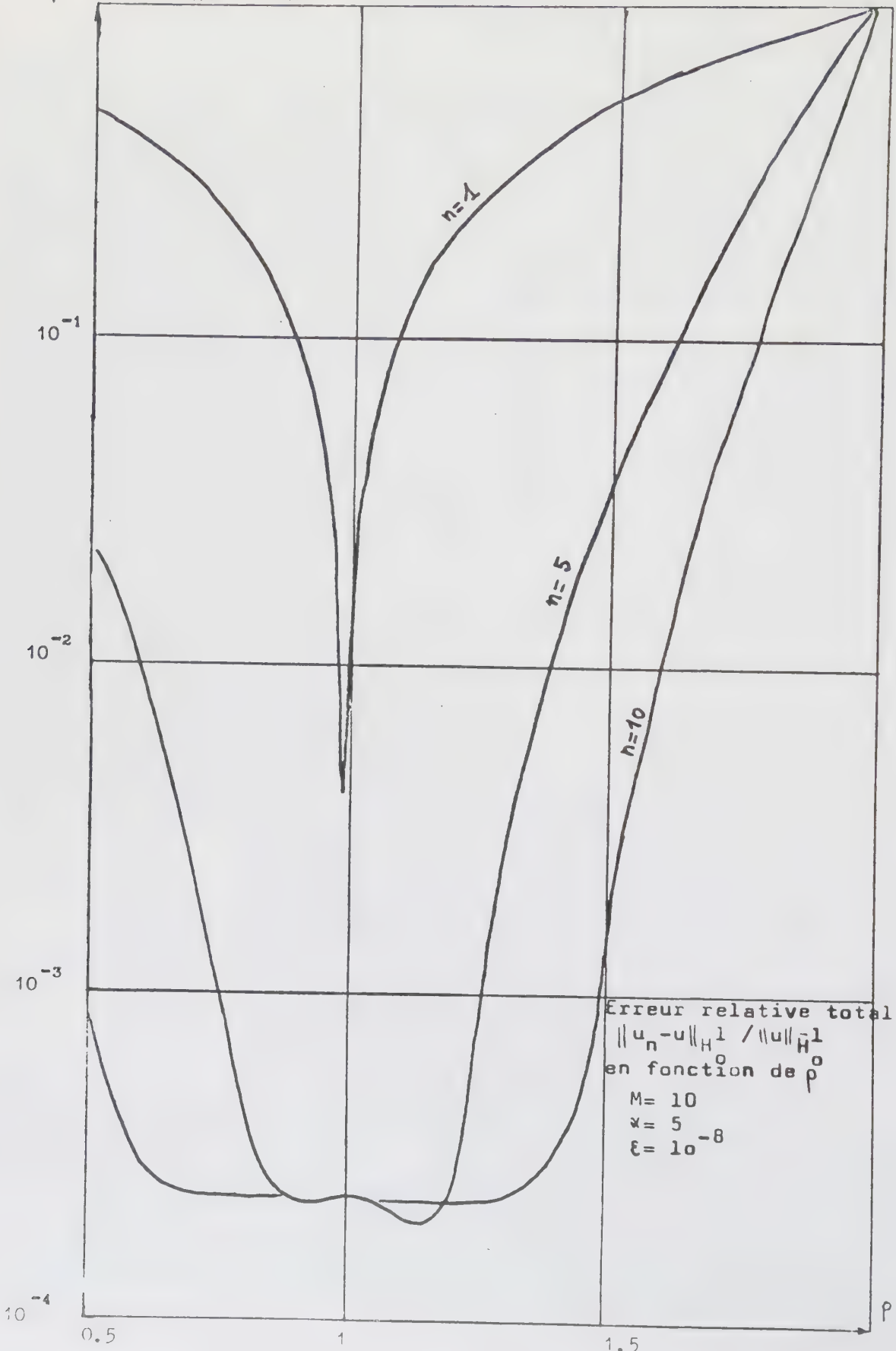
0.55	$3.0686 \cdot 10^{-1}$	$4.4284 \cdot 10^{-1}$	$1.7228 \cdot 10^{-2}$	$5.1265 \cdot 10^{-4}$	1080
0.60	2.8164	3.9219	$9.4836 \cdot 10^{-3}$	3.1408	984
0.67	2.4307	3.2129	3.6354	2.4414	869
0.74	2.0013	2.5038	1.0085	2.3434	769
0.8	$1.5925 \cdot 10^{-2}$	1.8961	$4.6716 \cdot 10^{-4}$	2.3352	691
0.84	1.2960	1.4910	3.0128	2.3348	640
0.9	$8.0996 \cdot 10^{-2}$	$8.8352 \cdot 10^{-2}$	2.3814	2.3348	542
0.92	6.3578	6.8117	2.3471	2.3349	535
0.94	4.5508	4.7903	2.3368	2.3349	505
0.96	2.6794	2.7758	2.3348	2.3350	468
0.97	1.7259	1.7785	2.3348	2.3350	445
0.98	$7.9530 \cdot 10^{-3}$	$8.2045 \cdot 10^{-3}$	2.3348	2.3350	412
0.982	6.2911	6.4920	2.3349	2.3351	402
0.983	5.5431	5.7132	2.3349	2.3351	395
0.984	4.8868	5.0188	2.3350	2.3352	387
0.985	4.3650	4.4485	2.3350	2.3352	375
0.986	4.0318	4.0548	2.3357	2.3356	335
0.987	3.9370	3.8918	2.3350	2.3352	337
0.988	4.0988	3.9879	2.3350	2.3352	381
0.989	4.4909	4.3259	2.3347	2.3350	393
0.990	5.0613	4.8555	2.3347	2.3350	402
0.995	9.1928	8.8698	2.3349	2.3351	423
1.00	$1.4048 \cdot 10^{-2}$	$1.3605 \cdot 10^{-2}$	2.3352	2.3353	431
1.02	3.4960	3.3524	2.3358	2.3358	458
1.05	6.8455	6.3807	2.3249	2.3350	529
1.1	$1.2954 \cdot 10^{-1}$	$1.1441 \cdot 10^{-1}$	2.1382	2.3358	577
1.2	2.7547	2.1568	2.5598	2.3374	712
1.3	4.6476	3.1697	$3.0083 \cdot 10^{-3}$	2.4369	907
1.4	7.2004	4.1827			
1.75	3.4093	7.7282	$2.7662 \cdot 10^{-1}$	$7.7161 \cdot 10^{-2}$	2264
1.85	6.9638	8.7412	5.1209	$2.6360 \cdot 10^{-1}$	
1.95	3.8615	9.7542	8.8592	$7.8801 \cdot 10^{-1}$	
ρ	$n = 2$	$n = 1$	$n = 5$	$n = 10$	nombre total d'itérations pour avoir une erreur stationnaire
	$\frac{\ u_n - u_{n-1}\ _{H_0^1}}{\ u_n\ _{H_0^1}}$		$\frac{\ u_n - u\ _{H_0^1}}{\ u\ _{H_0^1}}$		

Courbes n° 2d

Nombre d'itérations



1 $\|u_n - u\|_{H^0}^1 / \|u\|_{H^0}^1$ Courbes n° 2e



Erreur relative totale
 $\|u_n - u\|_{H^0}^1 / \|u\|_{H^0}^1$
 en fonction de ρ
 $M = 10$
 $\alpha = 5$
 $\xi = 10^{-8}$

Accélération de la convergence

Posons

$$G(u_h^n) = u_h^n - \rho S_h^{-1} (A_h u_h^n) , \quad u_h^0 = 0$$

$$l_n = -\rho S_h^{-1} (A_h u_h^n) = G(u_h^n) - u_h^n$$

On étudie alors la convergence de la suite u_h^n définie par (cf. chap. I remarque 1.2)

$$(2.22) \quad u_h^{n+1/2} = G(u_h^n) , \quad u_h^0 = 0$$

$$(2.23) \quad u_h^{n+1} = u_h^n - \frac{l_n^2}{l_{n+1/2} \cdot l_n}$$

Un seul résultat suffit à montrer l'énorme supériorité de cette deuxième méthode par rapport aux itérations (2.18).

Pour $M = 1/h = 10$, $\alpha = 5$, $\epsilon = 10^{-5}$, $\rho = 0.796$

On obtient le résultat :

$$\|u_h^n - u\| / \|u\| = 2,205 \cdot 10^{-4}$$

après 37 itérations par la 2ème méthode
et 77 itérations par la 1ère méthode

Ceci correspond à un temps de calcul identique si le phénomène suivant n'avait pas lieu :

Alors qu'à l'itération 2 ; $\|u_h^2 - u\| / \|u\| = 3,76 \cdot 10^{-2}$ par la première méthode ;

on a dès l'itération 1 ; $\|u_h^1 - u\| / \|u\| = 1,056 \cdot 10^{-4}$ par la deuxième méthode ! Ceci résulte aussi du bon choix de ϵ .

Je remercie M. J.F. Boutemy pour avoir effectué la programmation de ce problème.

II-3 Problème 3

Soit $\Omega = (0,1) \times (0,1)$ un ouvert de R^2 . On cherche

$$u \in W_0^{1,3}(\Omega) = \{u | u \in L^3(\Omega) , D_i u \in L^3(\Omega) \quad i = 1,2\}$$

vérifiant l'équation

$$(2.24) \quad Au = f \text{ pour } f \text{ donné dans } L^{3/2}(\Omega)$$

avec
$$Au = \sum_{i=1}^2 D_i (|D_i u| D_i u) - \lambda \Delta u, \quad \lambda > 0.$$

Au problème (2.24) on associe comme dans [3] et [14] l'équation discrétisée.

(2.25) $A_h u_h = f_h$; $h = (h_1, h_2)$ étant le paramètre de discrétisation d'un réseau $\mathcal{R}_h$ associé à l'ouvert Ω .

avec
$$A_h u_h = - \sum_{i=1}^2 \nabla_i (|\nabla_i u_h| \nabla_i u_h) - \lambda \sum_{i=1}^2 \nabla_i^2 u_h$$

où
$$\nabla_i u_h(x) = \frac{u_h(x + \frac{h_i}{2}) - u_h(x - \frac{h_i}{2})}{h_i} \quad i = 1, 2$$

Si $i = 1$ on a $u_h(x + \frac{h_i}{2}) = u_h(x_1 + \frac{h_1}{2}, x_2)$

De même nous avons

$$\nabla_i^2 u_h(x) = \frac{u_h(x+h_i) + u_h(x-h_i) - 2u_h(x)}{h_i^2} \quad i = 1, 2$$

Pour l'étude détaillée de la discrétisation du problème (2.24) cf. [3] et [14].
On se propose maintenant de résoudre le problème (2.25) à l'aide des méthodes itératives du chapitre I.

1ère méthode : méthode de décomposition de l'opérateur A_h .

Nous allons appliquer ici la méthode de la remarque 7.3 du chap. I de [14] en décomposant uniquement l'opérateur A_h cf [10'], [14'], [17].

Plus précisément posons :

(2.26) $A_h = A_h^1 + A_h^2$ et $f_h = f_h^1 + f_h^2$

avec
$$A_h^i u_h = - \nabla_i (|\nabla_i u_h| \nabla_i u_h) - \lambda \nabla_i^2 u_h \quad i = 1, 2.$$

Soient ϵ et k deux paramètres > 0 et N un entier tel que $kN = 1$. On suppose que $\epsilon \longrightarrow 0$ et $k = \frac{1}{N} \longrightarrow 0$ et $\frac{k}{\epsilon} \longrightarrow 0$.

On applique alors la méthode de décomposition donnée par les formules :

$$(2.27) \left\{ \begin{array}{l} A_h u_h^{1/2} + \frac{\epsilon}{k} u_h^{1/2} = f_1 + \frac{\epsilon}{k} u_h^{(0)} \quad (\text{on prend } u_h^{(0)} = 0) \\ A_h^2 u_h^{(1)} + \frac{\epsilon}{k} u_h^{(1)} = f_2 + \frac{\epsilon}{k} u_h^{(1/2)} \\ A_h^1 u_h^{(3/2)} + \frac{\epsilon}{k} u_h^{(3/2)} = f_1 + \frac{\epsilon}{k} u_h^{(1)} \\ A_h^2 u_h^{(2)} + \frac{\epsilon}{k} u_h^{(2)} = f_2 + \frac{\epsilon}{k} u_h^{(3/2)} \quad \text{etc...} \end{array} \right.$$

Comme les opérateurs $\left(A_h^1 + \frac{\epsilon}{k} I \right)$ et $\left(A_h^2 + \frac{\epsilon}{k} I \right)$

vérifient les hypothèses du théorème 1.1 chap. I sur les itérations ; on peut donc procéder comme suit : (on supprime l'indice h pour simplifier).

1) On se donne $u^{(0)} = 0$

2) On calcule $u^{(1/2)}$ donné par (2.27) à l'aide des itérations :

$$(2.28) \quad u_{n+1}^{(1/2)} = u_n^{(1/2)} - \rho_1 S_1^{-1} \left(A_1 u_n^{(1/2)} + \frac{\epsilon}{k} u_n^{(1/2)} - f_1 - \frac{\epsilon}{k} u^{(0)} \right)$$

$$\text{avec } S_1 = -\nabla_1^2 \quad \text{et } \rho_1 = \frac{\lambda}{\left(\frac{4}{\lambda \sqrt{h}} \|f_1 + \frac{\epsilon}{k} u^{(0)}\|_{L_2} + \lambda + \frac{\epsilon}{k} \right)^2} ; \quad u_0^{(1/2)} = 0$$

$$\text{on a : } \lim_{n \rightarrow \infty} u_n^{(1/2)} = u^{(1/2)}$$

3) Puis on calcule $u^{(1)}$ donné par (2.27) à l'aide des itérations :

$$(2.29) \quad u_{n+1}^{(1)} = u_n^{(1)} - \rho_2 S_2^{-1} \left(A_2 u_n^{(1)} + \frac{\epsilon}{k} u_n^{(1)} - f_2 - \frac{\epsilon}{k} u^{(1/2)} \right)$$

$$\text{avec } S_2 = -\nabla_2 ; \quad \rho_2 = \frac{\lambda}{\left(\frac{4}{\lambda \sqrt{h}} \|f_2\| + \|u^{(1/2)}\|_{L_2} + \lambda + \frac{k}{\epsilon} \right)^2}$$

$$\text{et } u_0^{(1)} = 0.$$

$$\text{On a alors } \lim_{n \rightarrow \infty} u_n^{(1)} = u^{(1)}$$

4) On sait alors que

$$\frac{1}{N} (u^{(0)} + u^{(1/2)} + u^{(3/2)} + \dots) \longrightarrow u_h$$

$$\text{et } \frac{1}{N} (u^{(1)} + u^{(2)} + u^{(3)} + \dots) \longrightarrow u_h$$

Si on suppose à priori que la solution exacte du problème (2.24) est donnée par la formule (en vue d'une expérience numérique) :

$$(2.30) \quad u = (x - x^2)(y - y^2) \quad (x, y) \in \Omega.$$

On en déduit, d'après (2.24) :

$$(2.31) \quad \begin{cases} f_1 = -D_1 (|D_1 u| D_1 u) - \lambda D_1^2 u \\ f_2 = -D_2 (|D_2 u| D_2 u) - \lambda D_2^2 u \end{cases}$$

Inversement on pose dans (2.27) f_1 et f_2 donnés par (2.31) et on calcule

$$(2.32) \quad \begin{cases} \frac{1}{N} (u^{(0)} + u^{(1/2)} + u^{(3/2)} + \dots) = v_h \\ \text{et } \frac{1}{N} (u^{(1)} + u^{(2)} + u^{(3)} + \dots) = w_h \end{cases}$$

Les fonctions v_h et w_h sont à comparer ensuite avec u solution de (2.24) donnée par (2.30).

Les résultats numériques sont les suivants :

1) Si on pose $\lambda = 20$; $h = \frac{1}{10}$; $\epsilon = 10^{-1}$; $k = 10^{-3}$; on a :

$$\frac{k}{\epsilon} = 10^{-2} \text{ et } N = \frac{1}{k} = 10^3 \text{ itérations à effectuer par les formules (2.27),}$$

Chacun des termes $u^{(p)}$ de cette sommation est obtenu comme limite d'une suite $u_n^{(p)}$ ($n \rightarrow \infty$) définie par les formules itératives (2.28) et (2.29).

Plus précisément le calcul de v_h et w_h donné par les formules (2.32) nécessite deux sommations, 10^3 fois l'inversion des opérateurs

$(A_h^1 + \frac{\epsilon}{h} I)$ et $(A_h^2 + \frac{\epsilon}{h} I)$ et dans chacune de ces inversions, il y a l'inversion d'un Laplacien en dimension une et à chaque itération. Le choix initial de $u_h^{(0)}$ et $u_0^{(p)}$ est toujours identiquement nul.

Nous obtenons alors

$$\|u - v_h\|_{L^2} = 8,307.10^{-5} \quad (h = \frac{1}{10})$$

$$\|u - w_h\|_{L^2} = 5,312.10^{-5} \quad (h = \frac{1}{10})$$

L'exécution et la compilation du programme entier en C.D.C. 3 600 ont duré 6 mn 54 s.

2) Pour $\lambda = 10$, $h = 10^{-1}$, $\epsilon = 10^{-1}$, $k = 10^{-3}$, $\frac{k}{\epsilon} = 10^{-2}$ et $N = 10^3$ nous obtenons les résultats suivants :

$$\|u - v_h\|_{L^2} = 7,7.10^{-4} \quad (h = 1/10)$$

$$\|u - w_h\|_{L^2} = 7,4.10^{-4} \quad (h = 1/10)$$

u étant la solution exacte du problème initial (2.24).

Temps de calcul = 1 mn 6 s (compilation + exécution) en C.D.C. 3 600.

Nous donnons maintenant pour $\lambda = 20$ et $\lambda = 10$; $h = 10^{-1}$ fixé; l'évolution des erreurs

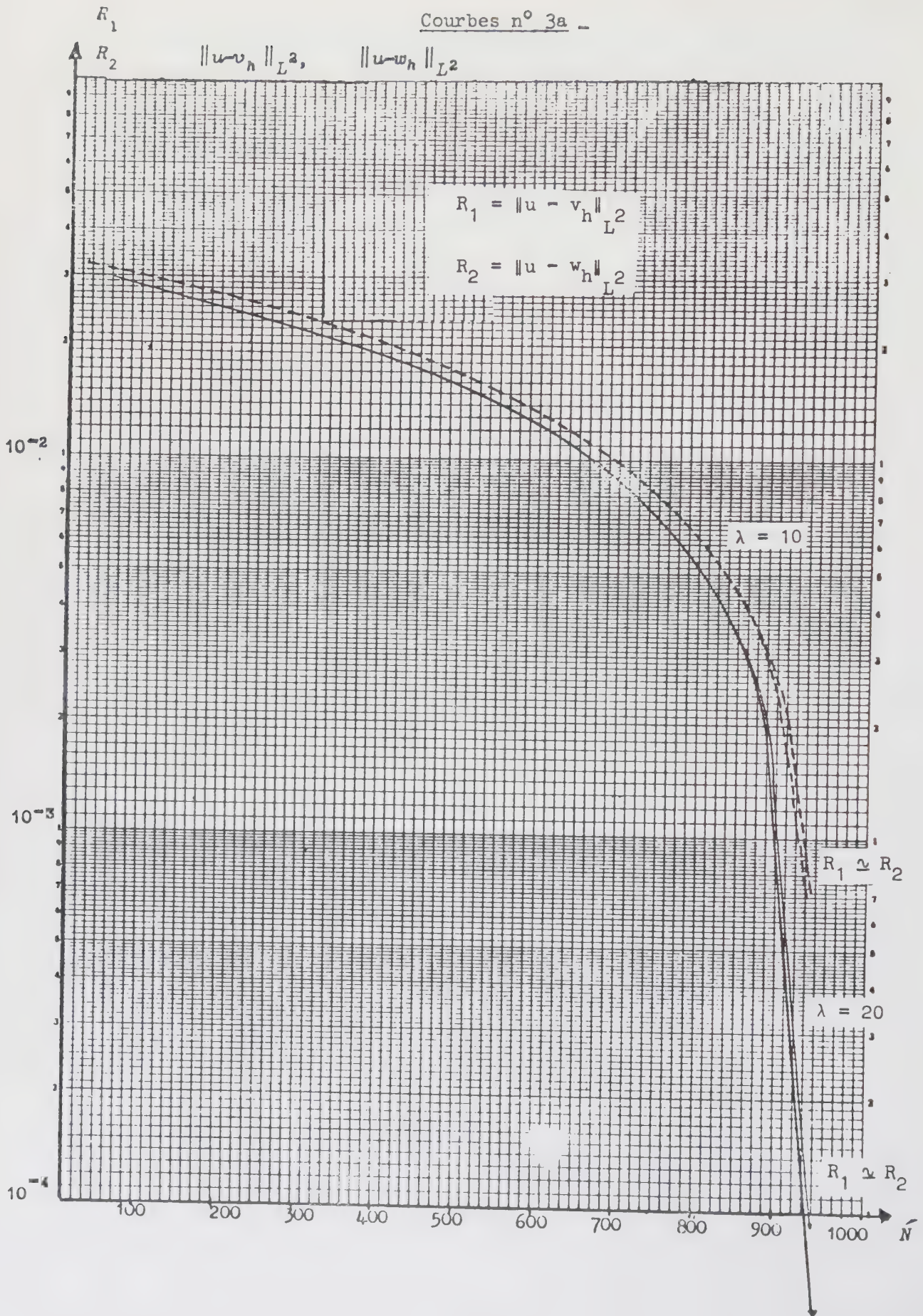
$R_1 = \|u - v_h\|_{L^2}$ et $R_2 = \|u - w_h\|_{L^2}$ en fonction de n itérations;

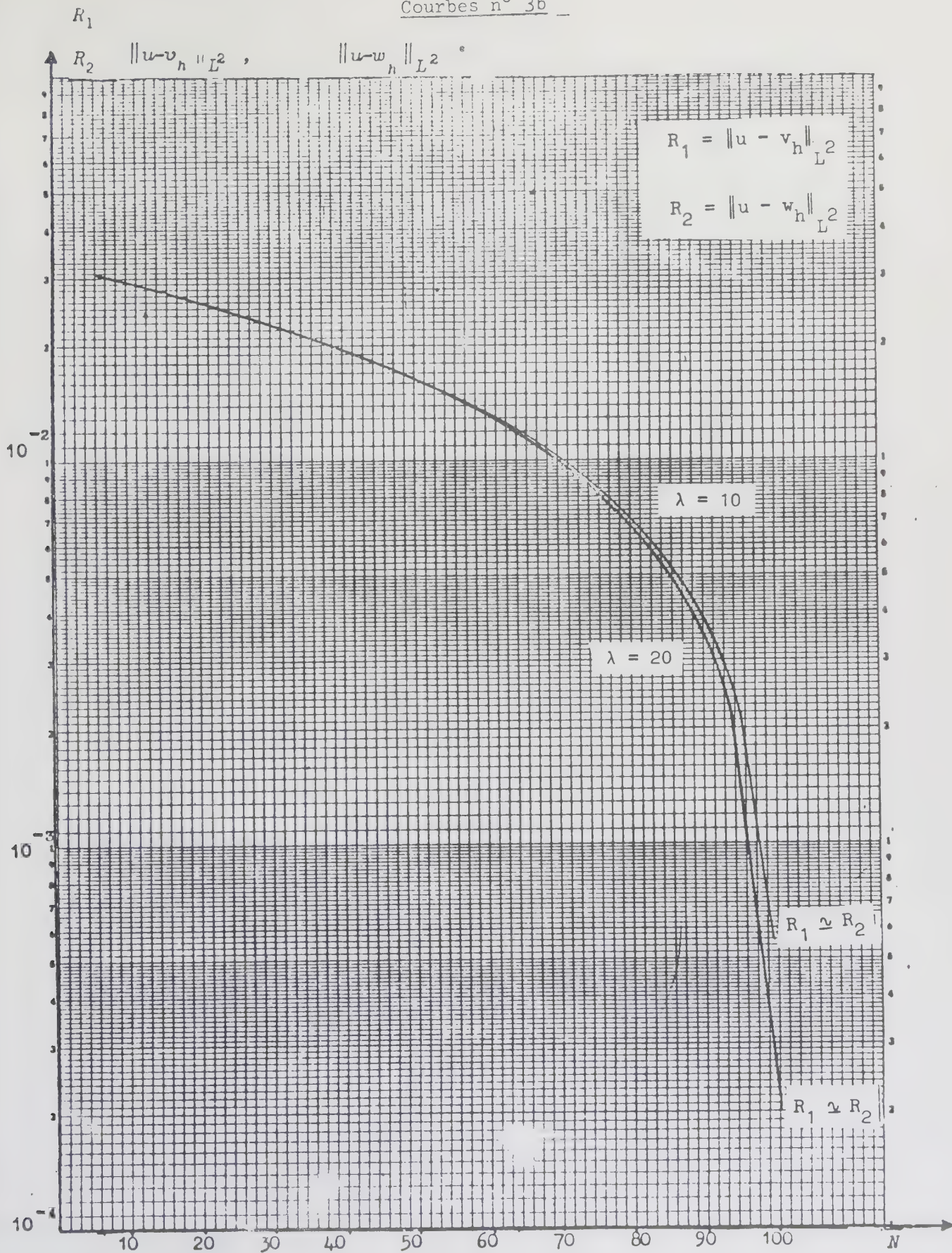
$$\text{où } v_h = \frac{1}{N} \sum_{n=0}^{N-1} u_h^{n+1/2} \quad \text{et} \quad w_h = \frac{1}{N} \sum_{n=0}^{N-1} u_h^n$$

Dans les courbes n° 3a on a fait : $\epsilon = 10^{-1}$, $k = 10^{-3}$, et $N = 1000$, ($h = 1/10$).

Dans les courbes n° 3b on a : $\epsilon = 10^{-1}$, $k = 10^{-2}$, et $N = 100$, ($h = 1/10$).

Courbes n° 3a -





Remarque 2.5

Nous avons montré que

$$v_h = \frac{1}{N} \sum_{n=0}^{N-1} u_h^{n+1/2} \longrightarrow u_h \text{ et } w_h = \frac{1}{N} \sum_{n=0}^{N-1} u_h^n \longrightarrow u_h$$

où u_h est la solution du problème (2.25)

On peut montrer les convergences suivantes :

$$(2.33) \quad \lim_{n \rightarrow 0} u_h^{n+1/2} = u_h \text{ et } \lim_{n \rightarrow \infty} u_h^n = u_h$$

Numériquement, la convergence (2.33) est réalisée pour $n \ll N$ ce qui est très avantageux.

Remarque 2.6

La méthode (2.26) de décomposition de l'opérateur A_h ramène donc un problème de dimension $n = 2$ à deux problèmes de dimension $n = 1$. Nous pouvons encore augmenter la compétitivité de la méthode en opérant comme suit :

A chaque étape, nous avons à résoudre les équations :

$$(2.34) \quad \left\{ \begin{array}{l} A_h^1 u_h^{n+1/2} + \frac{\epsilon}{k} u_h^{n+1/2} = f_1 + \frac{\epsilon}{k} u_h^n \\ \text{où } A_h^1 u = -\nabla_1 (|\nabla_1 u| \nabla_1 u) - \lambda \nabla_1^2 u \end{array} \right.$$

$$(2.35) \quad \left\{ \begin{array}{l} A_h^2 u_h^{n+1} + \frac{\epsilon}{k} u_h^{n+1} = f_2 + \frac{\epsilon}{k} u_h^{n+1/2} \\ \text{où } A_h^2 u = -\nabla_2 (|\nabla_2 u| \nabla_2 u) - \lambda \nabla_2^2 u \end{array} \right.$$

Au lieu de résoudre ces deux équations à l'aide des itérations internes (2.28), (2.29), on peut aussi remplacer (2.27), (2.32) par

$$(2.36) \quad -\nabla_1 (|\nabla_1 u_h^n| \nabla_1 u_h^{n+1/2}) - \lambda \nabla^2 u_h^{n+1/2} + \frac{\epsilon}{h} u_h^{n+1/2} = f_1 + \frac{\epsilon}{h} u_h^n$$

$$(2.37) \quad -\nabla_2 (|\nabla_2 u_h^{n+1/2}| \nabla_2 u_h^{n+1}) - \lambda \nabla^2 u_h^{n+1} + \frac{\epsilon}{h} u_h^{n+1} = f_2 + \frac{\epsilon}{h} u_h^{n+1/2}$$

Les deux équations (2.36), (2.37) se résolvent alors très rapidement par la méthode de Gauss.

Cette méthode semble générale.

2ème méthode : méthodes itératives directes

1) Cas où le paramètre ρ est fixe

On résout le problème (2.24) par la méthode itérative suivante :

$$(2.38) \quad S_h u_h^{n+1} = S_h u_h^n - \rho (A_h u_h^n - f)$$

avec
$$S_h = - \sum_{i=1}^2 \nabla_i^2 \quad \text{et} \quad \rho = \lambda / C^2(N)$$

avec
$$N = \frac{\sqrt{2}}{\lambda} \|f_h\|_{L^2} \quad \text{et} \quad C(N) = 2 \frac{N}{h} + \lambda, \quad u_h^0 = 0$$

On suppose à priori que la solution exacte s'écrit :

$$(2.39) \quad u = (x - x^2)(y - y^2)$$

On prendra alors $f_h' = f = -D_1 (|D_1 u| D_1 u) - D_2 (|D_2 u| D_2 u) - \lambda \Delta u$

Les itérations (2.38) permettent alors de calculer (u_h^n) pour h fixé, que l'on pourra comparer à u donné par (2.39).

A chaque itération principale (2.38) nous inversons l'opérateur

$$S_h = - \sum_{i=1}^2 \nabla_i^2 \quad \text{par la méthode bien connue de Gauss-Seidel cf [16]}$$

d'ailleurs moins rapide que la méthode de "relaxation successive" [16] utilisée dans la première méthode pour l'inversion de ce même S_h .

Dans les itérations internes de Gauss-Seidel on réinitialise chaque fois à 0. Le tableau suivant résume les résultats numériques de cette expérience pour $\lambda = 20$

h	$\ u - u_h^n\ _{L_2}$	Temps d'exécution en C.D.C.3 600
1/10	$6,34 \cdot 10^{-5}$	30,39 s pour 60 itérations
1/20	$2,43 \cdot 10^{-5}$	6mn 26 s pour 150 itérations

(cf les remarques de la page 61)

L'évolution de l'erreur $\|u - u_h^n\|_{L_2}$ pour $h = \frac{1}{10}$ et $h = \frac{1}{20}$ en fonction du nombre n d'itérations est résumée dans la courbe n° 3c

Dans les calculs précédents, nous avons posé

$$(2.40) \quad \nabla_1^2 u(i,j) = \frac{u(i+2,j) + u(i-2,j) - 2u(i,j)}{4h^2}$$

et une formule analogue pour $\nabla_2^2 u(i,j)$

Quand on pose :

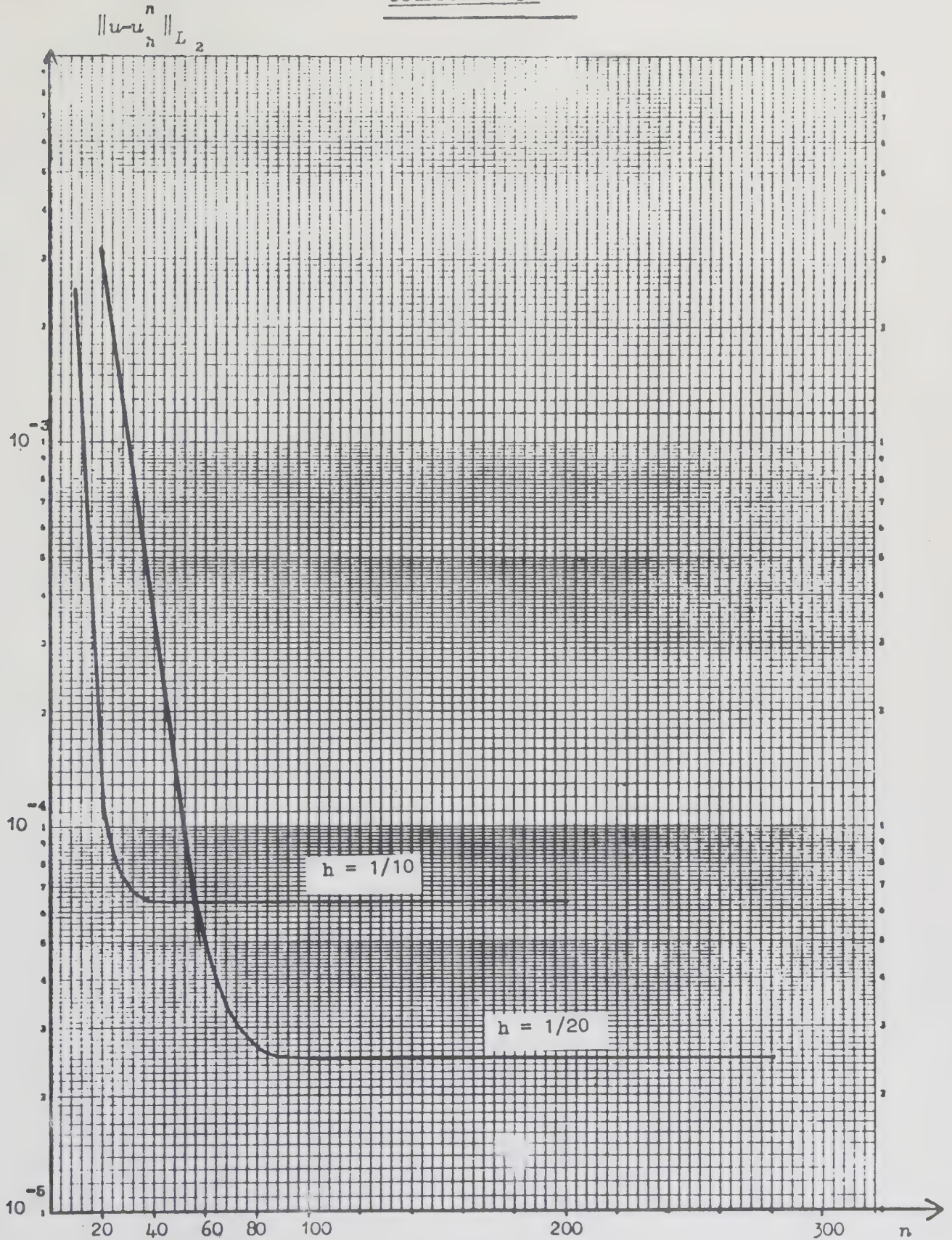
$$(2.41) \quad \hat{\nabla}_1^2 u(i,j) = \frac{u(i+1,j) + u(i-1,j) - 2u(i,j)}{h^2}$$

et une formule analogue pour $\hat{\nabla}_2^2 u(i,j)$; nous obtenons les courbes n° 3d et bien entendu les limites de convergences sont exactement les mêmes ; mais le temps d'exécution est approximativement divisé par deux.

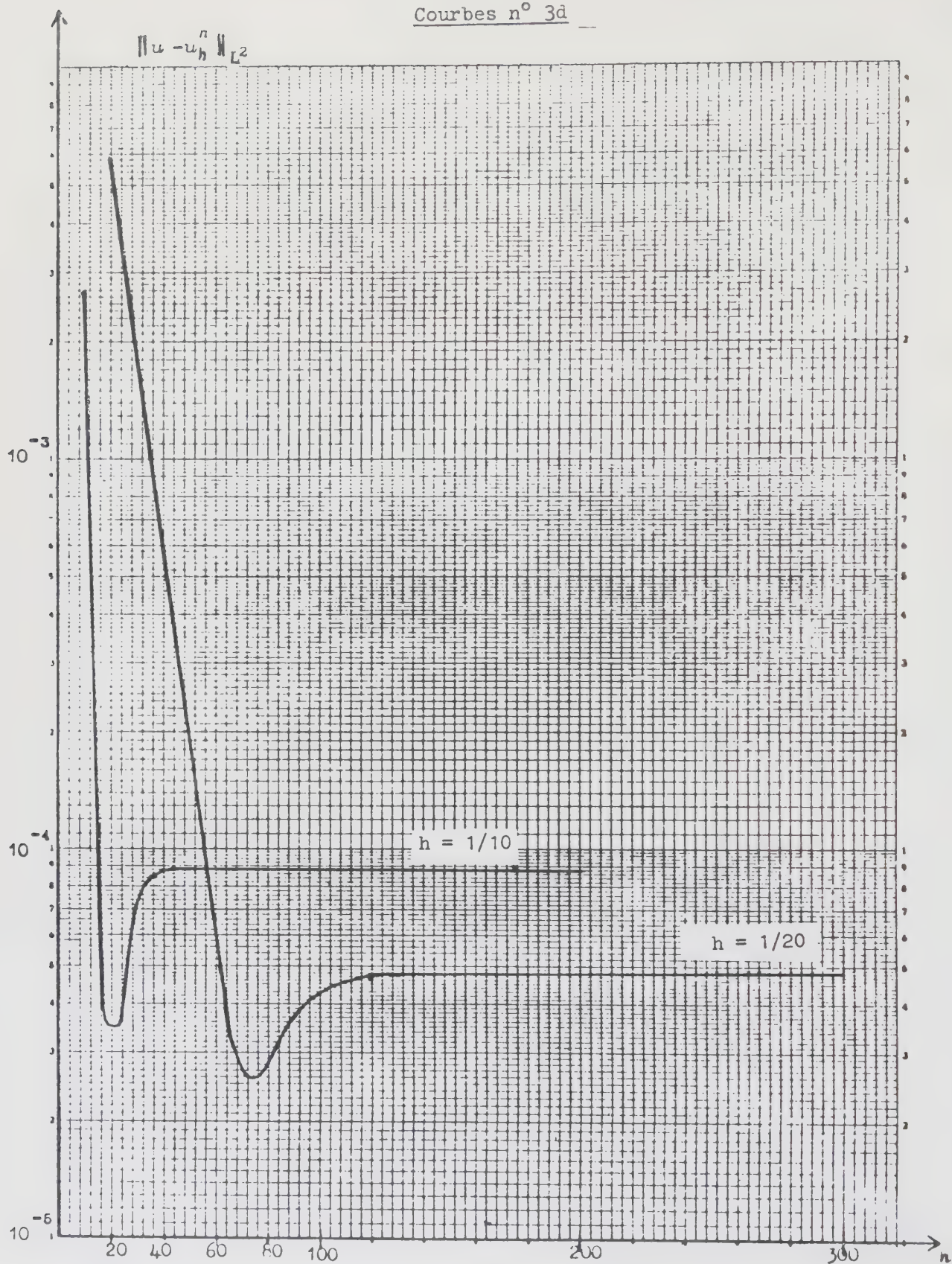
Pour $h = 1/10$ on effectue 60 itérations en 11 s.

Pour $h = 1/20$ on effectue 150 itérations en 3 mn 28 s.

Courbes n° 3c



Courbes n° 3d



2) Quelques remarques

1- L'équation (2.38) peut encore s'écrire

$$(2.42) \quad u_h^{n+1} = S_h^{-1} (S_h u_h^n - P (A_h u_h^n - f_h))$$

Comme on inverse l'opérateur S_h par la méthode itérative de Gauss-Seidel, on peut prendre comme valeur initiale de u_h^{n+1} pour $m = 0$ le vecteur u_h^n de l'itération précédente. Ceci améliore grandement la convergence (le Laplacien est alors inversé en quelques itérations seulement) et le temps d'exécution du programme.

2- L'équation (2.42) peut encore s'écrire

$$(2.43) \quad u_h^{n+1} = u_h^n - P S_h^{-1} (A_h u_h^n - f_h)$$

Sous cette forme, les erreurs d'arrondi interviennent moins dans l'exécution des calculs.

Compte tenu de la remarque précédente, et si de plus on fait $f_h = A_h u$ dans (2.43) on obtient pour $\lambda = 10$, $h = 1/15$ au bout de $n = 500$ itérations $\|u_h^n - u\|_{L^2} = 7.10^{-12}$ soit à peu de chose près la solution exacte. Les calculs ont duré (compilation + exécution) 2mn en C.D.C. 3600.

3- Si on fait $\rho_0 = \lambda / C(N)$ avec $C(N) = 2 \frac{N}{h} + \lambda$ et $N = \left(\frac{1}{2^{1/3}} \|f\|_{L_{3/2}} \right)^{1/2}$ la convergence est plus lente. On peut la rendre plus rapide en surestimant le paramètre ρ .

Les courbes n° 3e donnent l'évolution de l'erreur $\|u_h^n - u\|_{L_2}$ en fonction du nombre d'itérations n pour $h = 1/10$, et pour $\rho = \rho_0$, $\rho = 40 \rho_0$, $\rho = 50 \rho_0$ et $\rho = 60 \rho_0$ (on a fait $f_h = f$).

Nous avons alors les résultats suivants : ($h = 1/10$, $\lambda = 10$)

On pose

$$E = \|u_h^n - u\|_{L_2}$$

N = nombre d'itérations à partir duquel on obtient la stabilisation des résultats

T = temps de calcul (compilation + exécution) en C.D.C. 3 600.

	ρ	$40 \rho_0$	$50 \rho_0$	$60 \rho_0$
E	$1,12 \cdot 10^{-5}$	$1,12 \cdot 10^{-5}$	$1,12 \cdot 10^{-5}$	$1,12 \cdot 10^{-5}$
N	480	80	100	80
T	1 mn 09 s pour 500 itérations	13 s pour 100 itérations	15 s pour 100 itérations	14 s pour 100 itérations

3) Cas où le paramètre ρ est variable avec les itérations n .

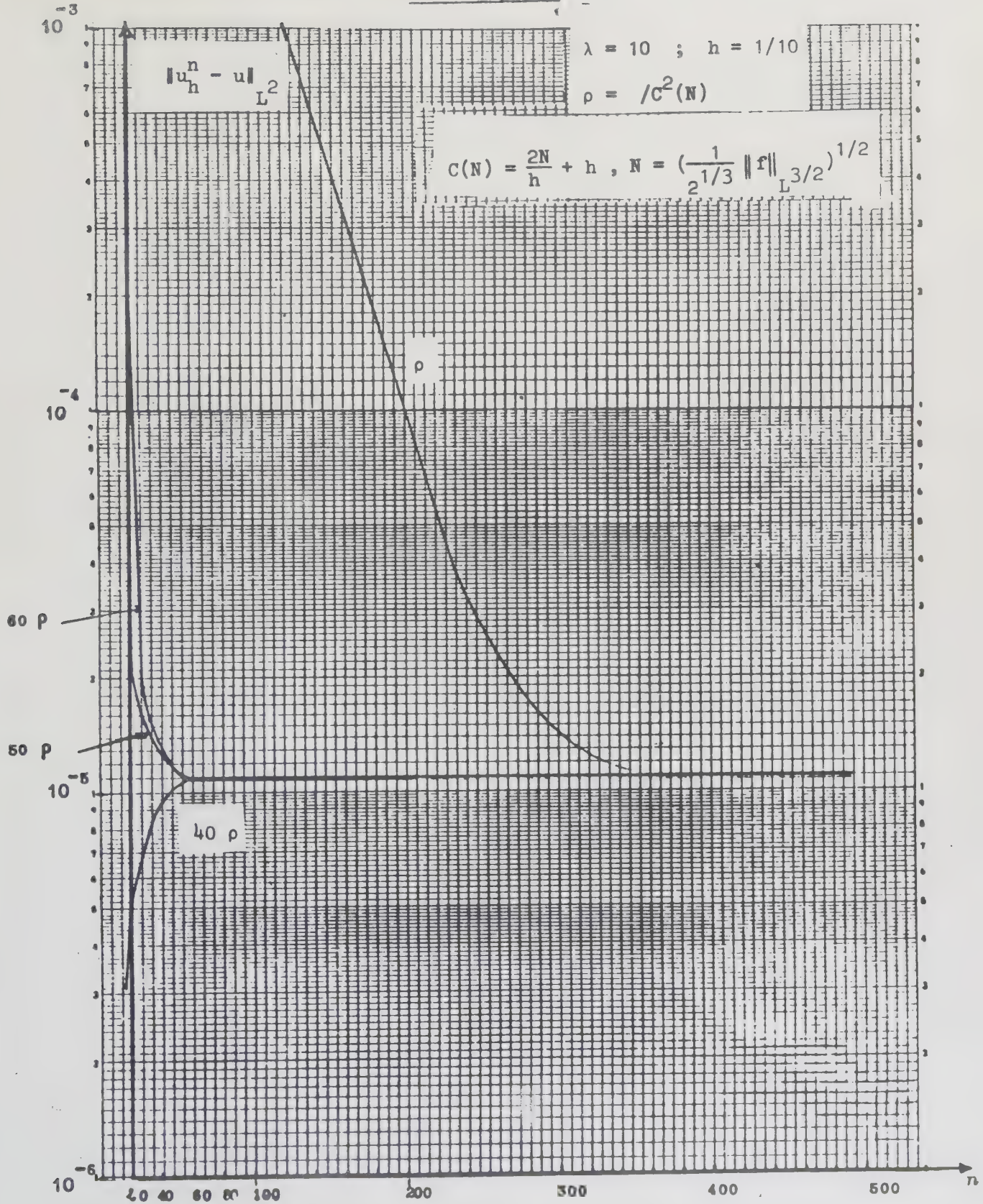
Les itérations (2.43) peuvent aussi être effectuées avec un paramètre ρ variable donné par

$$\rho_n = \frac{(Au_n - Au_{n-1}, u_n - u_{n-1})}{[S^{-1}(Au_n - Au_{n-1})]^2}$$

Nous avons fait l'expérience pour $h = 1/10$, $\lambda = 10$, $f_h = A_h u$
Au bout de $n = 400$ itérations on a : $\|u_h^n - u\|_{L_2} = 3,59 \cdot 10^{-12}$ soit
pratiquement une erreur relative de l'ordre de 10^{-10} .

Le temps de calcul (compilation + exécution) en C.D.C. 3 600 est de
2 mn 6 s pour 400 itérations.

Courbes n° 3e



2- Problème 2 : Régularisation elliptique

On cherche $u \in W_0^{1,3}(\Omega)$ vérifiant l'équation

$$(2.44) \quad Bu = f \quad \text{pour } f \text{ donné dans } L^{3/2}(\Omega)$$

$$\text{avec} \quad Bu = - \sum_{i=1}^2 D_i (|D_i u| D_i u)$$

Au problème (2.44) on associe alors le problème (cf. théorème 4.1 de [14])

$$(2.45) \quad Au = f$$

$$\text{avec} \quad Au = - \sum_{i=1}^2 D_i (|D_i u| D_i u) - \lambda \Delta u = Bu - \lambda \Delta u$$

avec $\lambda > 0$ destiné à tendre vers 0.

Nous reprenons alors le problème discrétisé (2.25) sous la forme

$$(2.46) \quad A_h u_h = f_h = f$$

et les itérations

$$(2.47) \quad u_h^{n+1} = u_h^n - \rho_n S_h^{-1} (A_h u_h^n - f_h)$$

$$\begin{aligned} \text{avec} \quad S_h &= - \sum_{i=1}^2 \nabla_i^2 \\ A_h u_h^n &= - \sum_{i=1}^2 \nabla_i (|\nabla_i u_h^n| \nabla_i u_h^n) - \lambda \sum_{i=1}^2 \nabla_i^2 u_h^n \\ \rho_n &= \frac{(A_h u_h^n - A_h u_h^{n-1}, u_h^n - u_h^{n-1})}{[S_h^{-1} (A_h u_h^n - A_h u_h^{n-1})]^2} \end{aligned}$$

On sait alors qu'on a les convergences suivantes : (cf. théorème 4.1 de [14])

$$u_h^n \xrightarrow{n \rightarrow \infty} u_h, \quad u_h \text{ étant la solution de } (2.46)$$

$$p_h u_h \xrightarrow{h \rightarrow \infty} u_\lambda, \quad u_\lambda \text{ étant la solution de } (2.45)$$

Enfin $u_\lambda \xrightarrow{\lambda \rightarrow \infty} u$, u étant la solution du problème (2.44).

Pour approcher la solution u du problème (2.44) nous allons résoudre (2.47) pour n suffisamment grand, $\lambda = 1/50$ et $h = 1/10$.

On se donne une solution exacte du problème (2.44) :

$$(2.48) \quad u = (x - x^2)(y - y^2)$$

on pose alors

$$f_h = f = B u$$

On calcule $(u_{h,\lambda}^n)$ à l'aide des itérations (2.47) et on compare le résultat avec u donné par (2.48). Nous obtenons au bout de $n = 360$ itérations l'erreur suivante :

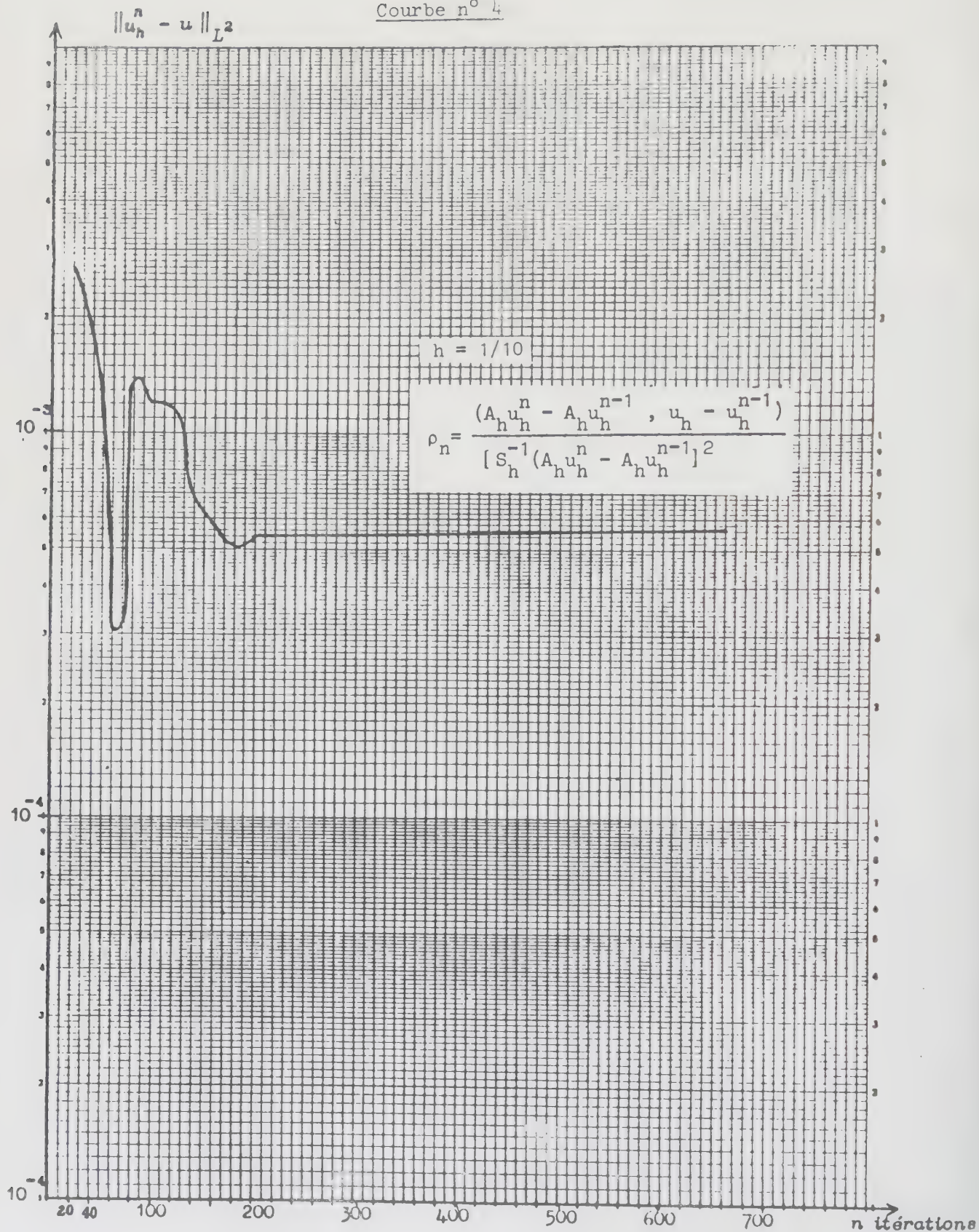
$$\|u_{h,\lambda}^n - u\|_{L_2} = 5,289.10^{-8}$$

pour $N \geq 360$ les résultats sont stationnaires.

Le calcul (compilation + exécution) en C.D.C. 3 600 a duré
2 mn 12 s pour $n = 500$ itérations.

Dans la courbe n° 4 nous donnons pour $h = 1/10$ l'évolution de l'erreur $\|u_h^n - u\|_{L_2}$ en fonction des itérations n .

Courbe n° 4



II-5 Problème 5

Soit $\Omega =]0,1[\times]0,1[$

On cherche $u \in W_0^{1,3}(\Omega)$ vérifiant l'équation :

$$(2.49) \quad -\Delta u + |u|u = f \quad \text{où } f \in L^{2/3}(\Omega).$$

Le problème discrétisé s'écrit dans l'espace V_h . (cf. [14])

$$(2.50) \quad -\sum_{i=1}^n \nabla_i^2 u_h + |u_h|u_h = f_h$$

La suite (u_h^n) définie par l'itération

$$(2.51) \quad u_h^{n+1} = u_h^n - \rho S_n^{-1} (A_h u_h^n - f_h)$$

où
$$A_h u_h = -\sum_{i=1}^n \nabla_i^2 u_h + |u_h|u_h$$

$$S_h = -\sum_{i=1}^n \nabla_i^2$$

$$(2.52) \quad \rho = 1 / \left(1 + \frac{\|f_h\|}{n} \right)^2$$

Converge vers u_h solution de (2.50) et on a $p_h u_h \xrightarrow[n \rightarrow 0]{} u$ solution de (2.49) ; où p_h est l'opérateur de prolongement de V_h dans $V = W_0^{1,3}(\Omega)$.

Nous savons alors qu'on a la relation : (cf. théorème 1.1 chap I)

$$(2.53) \quad \|u_h^{n+1} - u_h^n\|^2 \leq \theta \|u_h^n - u_h^{n-1}\|^2 \quad \text{avec}$$

$$(2.54) \quad \theta = 1 - \rho$$

θ est alors le coefficient de contraction des itérations (2.51).

Nous voyons alors que

pour $h = 1/10$ $\rho = 1,510^{-2}$ et $\theta = 0,985$

pour $h = 1/20$ $\rho = 4,3410^{-3}$ et $\theta = 0,995$

la convergence est donc très mauvaise pour $h = 1/10$ (domaine de 100 points) et impossible matériellement pour $h = 1/20$ (réseau à 400 points). Pour la dimension $n = 1$ cette méthode était encore suffisante

Une nouvelle estimation du paramètre ρ .

Dans la formule (2.52) on peut encore écrire

$$(2.55) \quad \rho = 1/C^2(N) \quad \text{avec} \quad C(N) = 1 + \frac{2N}{2^{3/2}h} \quad \text{et}$$

$$(2.56) \quad N = \sqrt{2} \|f_h\|$$

Par ailleurs on sait que (cf. Théorème 1.1 chap I)

$$(2.57) \quad N \geq 2 [u_h] = 2\sqrt{2} \|u_h\|_{L_2}$$

Le meilleur choix de la constante N est bien entendu $2[u_h] = 2\sqrt{2} \|u_h\|_{L_2}$ où u_h est la solution inconnue de (2.49). L'estimation (2.56) est trop large par rapport à l'estimation

$$(2.58) \quad N_0 = 2 [u_h]$$

Mais

$$(2.59) \quad [u_h] = [u_h - u_h^n + u_h^n] \leq [u_h - u_h^n] + [u_h^n]$$

Par ailleurs

$$[u_h^n - u_h]^2 \leq (A_h u_h^n - A_h u_h, u_h^n - u_h) \leq \|A_h u_h - A_h u_h^n\|^* \|u_h - u_h^n\|$$

Mais

$$\|u_h - u_h^n\| = \frac{[u_h - u_h^n]}{\sqrt{2}} \quad \text{et} \quad A_h u_h = f_h$$

Donc

$$[u_h - u_h^n]^2 \leq \frac{1}{\sqrt{2}} \|f_h - A_h u_h^n\|^* [u_h - u_h^n]$$

Soit

$$(2.60) \quad [u_h - u_h^n] \leq \frac{\|f_h - A_h u_h^n\|^*}{\sqrt{2}}$$

Les formules (2.59) et (2.60) donnent :

$$[u_h] \leq \frac{1}{\sqrt{2}} \|f_h - Au_h^n\| + \sqrt{2} \|u_h^n\|$$

On pose alors

$$(2.61) \quad N_n = \sqrt{2} \|f_h - A_h u_h^n\| + 2\sqrt{2} \|u_h^n\|$$

Et d'après (2.55) on a :

$$(2.62) \quad \rho = \rho_n = 1 / \left(1 + \frac{\|f_h - A_h u_h^n\| + 2 \|u_h^n\|}{h} \right)^2$$

On prendra $u_h^0 = 0$

Remarque 2.7

Dans (2.61) on voit que $\lim_{n \rightarrow \infty} N_n = N_0 = 2[u_h]$ qui est la meilleure estimation possible.

Remarque 2.8

On peut donc effectuer les itérations (2.51) de plusieurs manières :

- a) utiliser ρ fixe donné par (2.52); ce qui est encore possible quand $h = 1/10$.
- b) utiliser $\rho = \rho_n$ donné par (2.62).
- c) utiliser $\rho = \rho_n$ jusqu'à $n = p$ itérations et continuer avec $\rho = \rho_p$ fixe.
- d) on peut aussi faire p itérations avec ρ fixe donné par (2.52) et poursuivre avec (ou reprendre les calculs avec)

$$(2.63) \quad \rho = \rho_p = 1/C^2(N_p) \quad \text{avec} \quad C(N_p) = 1 + \frac{2N_p}{2^{3/2}h} \quad \text{et} \quad N_p = 2[u_h^n]$$

Nous avons essayé toutes ces méthodes. Nous notons par R_a, R_b, R_c, R_d les résultats obtenus par les procédés a), b), c), d) successivement.

On se donne une solution exacte de (2.49)

$$(2.64) \quad u(x, y) = (x - x^2)(y - y^2)$$

Ce qui donne

$$(2.65) \quad f = 2[(x - x^2) + (y - y^2)] + (x - x^2)^2 + (y - y^2)^2$$

on pose dans (2.51) $f_h = f$ et on calcule u_h^n à l'aide de (2.51) que l'on compare à u donné par (2.64).

Nous avons les résultats suivants pour $h = 1/10$
(domaine de 100 points intérieurs) :

	Nombre d'itérations	$\ u_h^n - u\ _{L_2}$	Exécution + compilation en C.D.C.3600
Résultats R_a $h = 1/10$	250 la convergence se poursuit	$7,48 \cdot 10^{-4}$	9 mn
Résultats R_b $h = 1/10$	90 après résultats stationnaires	$1,59 \cdot 10^{-11}$	3 mn 18 s
Résultats R_d $p = 40$ $h = 1/10$	80 après résultats stationnaires.	$9,80 \cdot 10^{-12}$	4 mn pour 40 + 80 itérations

Remarque 2.9

Dans la méthode b) de la remarque précédente, le coefficient de contraction des itérations (2.51) est donné par la relation

$$\theta = 1 - \rho_n$$

pour $n = 10$; $\theta_{10} = 0,981$

pour $n = 10$; $\theta_{90} = 0,640$

II-6 Problème 6

Soit Ω un ouvert borné de $\mathbb{R}^n$ suffisamment régulier, On munit l'espace $W_0^{1,p}(\Omega)$ de la norme $\|u\| = \|\nabla u\|_{L^p(\Omega)}$ avec

$$\|\nabla u\| = \left(\sum_{i=1}^n |D_i u|^2 \right)^{1/2}, \text{ qui en fait un Banach uniformément convexe.}$$

On cherche $u \in W_0^{1,p}(\Omega)$ vérifiant

$$(2.66) \quad A u = f \text{ pour } f \text{ donné dans } W^{-1,q}(\Omega), \quad \frac{1}{p} + \frac{1}{q} = 1;$$

$$\text{avec } A u = \sum_{i=1}^n D_i (|\nabla u|^{p-2} D_i u)$$

on sait (cf. [14] chap. II, § II-3) que le problème (1.66) admet une solution unique.

Nous lui associons alors comme dans [14] l'équation discrétisée avec les notations habituelles :

$$(2.67) \quad A_h u_h = f_h \quad u_h \in V_h$$

$$\text{avec } A_h u_h = - \sum_{i=1}^n \nabla_i (|\nabla_h u_h|^{p-2} \nabla_i u_h) \quad \text{où } \nabla_h u_h = \left(\sum_{i=1}^n |\nabla_i u_h|^2 \right)^{1/2}$$

$$f_h = \frac{1}{h_1 \dots h_n} \int_{\Omega} f \theta_h^M dx$$

Nous allons maintenant résoudre l'équation (2.67) à l'aide d'une méthode itérative en nous plaçant dans le cadre du théorème 3.1 chap. I.

$$\text{Posons } S_h u_h = -\Delta_h u_h = - \sum_{i=1}^n \nabla_i^2 u_h.$$

On munit V_h du produit scalaire $(u_h, v_h)_h = (u_h, v_h)_{L^2}$

et d'une nouvelle norme $\|u_h\|_h^2 = \sum_{i=1}^n \|\nabla_i u_h\|_{L^2}^2$ équivalente à

$$\|u_h\|_h^2 = (u_h, u_h)_h. \text{ Soit } B_h u_h = \Delta_h u_h. \text{ Nous avons alors la}$$

Proposition 2.3

La suite $(u_{h,r}^{n+1})$ définie par l'itération $(u_{h,r}^0 = 0)$

$$(2.68) \quad u_{h,r}^{n+1} = u_{h,r}^n - \rho S_h^{-1} (A_h u_{h,r}^n + \epsilon_r \Delta_h u_{h,r}^n - f_h - \epsilon_r g_h)$$

avec g_h donné dans V_h ; converge vers la solution u_h de l'équation (2.66), quand $n, r \rightarrow \infty$ (avec $\epsilon_r \rightarrow 0$ quand $r \rightarrow \infty$).

Démonstration

Les hypothèses H5), H6) du paragraphe I-3. chap. I sont trivialement vérifiées. Nous allons vérifier H7) en calculant explicitement les paramètres k , N , $C(N)$ et enfin ρ .

Posons $L_h u_{h,r} = A_h u_{h,r} - \epsilon_r \Delta_h u_{h,r}$ Nous avons alors

$$a) \quad (L_h u_{h,r} - L_h v_{h,r}, u_{h,r} - v_{h,r}) \geq \epsilon_r [u_{h,r} - v_{h,r}]$$

ce qui donne $k = \epsilon_r$

b) Supposons que $[u_{h,r}] \leq N$ et $[v_{h,r}] \leq N$; il vient :

$$[u_{h,r}]^2 = \sum_{i=1}^n \|\nabla_i u_{h,r}\|_L^2 = \sum_{i=1}^n h_1 \dots h_n \sum_{M \in \Omega'_h} |\nabla_i u_{h,r}(M)|^2$$

$$\text{Donc} \quad |\nabla_i u_{h,r}(x)| \leq \frac{[u_{h,r}]}{(h_1 \dots h_n)} = N' \quad \text{Par ailleurs}$$

$$(2.69) \quad |(A_h u_{h,r} - A_h v_{h,r}, w_{h,r})_h| \leq \left(\sum_{i=1}^n \|\nabla_i u_{h,r}\|^{p-2} \|\nabla_i u_{h,r} - \nabla_i v_{h,r}\|^{p-2} \|\nabla_i v_{h,r}\|^2 \right)^{1/2} \\ \times \left(\sum_{i=1}^n \|\nabla_i w_{h,r}\|^2 \right)^{1/2}$$

On montre facilement que si

$$|\alpha_i| \leq N' \quad , \quad |\beta_i| \leq N' \quad ; \quad \alpha_i, \beta_i \text{ réels} \quad i = 1, \dots, n.$$

On a l'inégalité :

$$(2.70) \quad \left| \left(\sum_{i=1}^n \alpha_i^2 \right)^{\frac{p-2}{2}} \alpha_i - \left(\sum_{i=1}^n \beta_i^2 \right)^{\frac{p-2}{2}} \beta_i \right| \leq (p-1) n^{\frac{p-2}{2}} N^{p-2} \sum_{i=1}^n |\alpha_i - \beta_i|$$

En posant $\alpha_i = \nabla_i u_{h,r}(x)$; $\beta_i = \nabla_i v_{h,r}(x)$, (2.69) devient :

$$(2.71) \quad |(A_h u_{h,r} - A_h v_{h,r}, w_{h,r})_h| \leq (p-1) n^{\frac{p-2}{2}} N^{p-2} [u_{h,r} - v_{h,r}] [w_{h,r}] .$$

Ce qui donne :

$$(2.72) \quad (L_h u_{h,r} - L_h v_{h,r}, w_{h,r}) \leq C(N) [u_{h,r} - v_{h,r}] [w_{h,r}] \quad \text{avec}$$

$$(2.73) \quad C(N) = \frac{(p-1) n^{\frac{p-2}{2}}}{(h \dots h)_{i=1}^{\frac{p-2}{2}}} + \epsilon_r .$$

Calcul de N

Nous avons successivement

$$|\nabla_i u_h| \leq \sqrt{n} \left(\sum_{i=1}^n |\nabla_i u_h|^2 \right)^{1/2} \quad i = 1, \dots, n .$$

$$\sum_{i=1}^n \|\nabla_i u_h\|_{L_p}^p \leq n^{\frac{p+2}{2}} \|\nabla_h u_h\|_{L_p}^p = n^{\frac{p+2}{2}} (A_h u_h, u_h)_h$$

D'après une inégalité de Poincaré :

$$\|u_h\|_{L_p} \leq C_1(\Omega) \left(\sum_{i=1}^n \|\nabla_i u_h\|_{L_p}^p \right)^{1/p} \quad \text{où } C_1(\Omega) \text{ est une constante}$$

ne dépendant que de l'ouvert Ω . On en déduit que

$$\left(\sum_{i=1}^n \|\nabla_i u_h\|_{L_p}^p \right)^{1/p} \leq \left(C_1(\Omega) n^{\frac{p+2}{2}} \|f_h + \epsilon_r g_h\|_{L_q} \right)^{q/p}$$

or $L_p \subset L_2$ avec injection continue pour $p \geq 2$. Donc

$$[u_h] = \left(\sum_{i=1}^n \|\nabla_i u_h\|_{L^2}^2 \right)^{1/2} \leq \left(C_1(\Omega) n^{\frac{p+2}{2}} \|f_h + \epsilon_r g_h\|_{L_q} \right)^{q/p}.$$

D'où

$$(2.74), \quad N = \left(C_1(\Omega) n^{\frac{p+2}{2}} \|f_h + \epsilon_r g_h\|_{L_q} \right)^{q/p}. \text{ Enfin en posant } u_{h,r}^0 = 0 \text{ et } \rho = \epsilon_r / C^2(N) \text{ nous avons la proposition.}$$

Remarque 2.10

$$1 - \quad \Omega = (0,1) \text{ nous avons } N = \|f_h + \epsilon_r g_h\|_{L_q}^{q/p} \text{ et}$$

$$C(N) = (p-1) \frac{N^{\frac{p-2}{2}}}{h^{\frac{p-2}{2}}} + \epsilon_r \quad \text{enfin} \quad \rho = \epsilon_r / C^2(N)$$

2 - Nous sommes aussi dans les hypothèses du théorème 3.3 chap. I. On a les itérations :

$$A_h u_h^n + \frac{\epsilon}{k} u_h^n = f_h + \frac{\epsilon}{k} u_h^{n-1}; \quad u_h^0 = 0$$

$$\text{et } \frac{1}{N} \sum_{n=1}^N u_h^n \longrightarrow u_h \text{ solution de (2.67);}$$

où $\epsilon \longrightarrow 0$, $k \longrightarrow 0$ ($\epsilon, k > 0$) avec $\frac{k}{\epsilon} > 0$ et N est un entier tel que $kN = 1$.

3 - Il peut être avantageux de prendre un paramètre $\rho = \rho_n$ variable avec les itérations. En particulier on peut poser $\epsilon_r = \epsilon$ et $\rho_n = \rho$ donné par (2.74) jusqu'à $n = n_0$ et pour $n \geq n_0$ on pose :

$$\rho_{n+1} = \epsilon / C^2(N'_n) \quad \text{avec} \\ C(N'_n) = (p-1) n^{\frac{p-2}{2}} N_n^{p-2} + \epsilon$$

$$\text{et } N'_n = \sup_i |\nabla_i u_h^n|$$

ou encore

$$\rho_{n+1} = \epsilon / C^2(N_n) \quad \text{avec} \\ C(N_n) = \frac{(p-1) n^{\frac{p-2}{2}} N_n^{p-2}}{(h_1 \dots h_n)^{\frac{p-2}{2}}} + \epsilon$$

$$\text{et } N_n = [u_n]$$

III - INEQUATIONS AUX DERIVEES PARTIELLES NON LINEAIRES ET PROBLEME D'OPTIMISATION SUR UN CONVEXE FERME.

Soit Ω un ouvert borné de R^n de frontière $\Gamma = \partial \Omega$ suffisamment régulière. On note par $x = (x_1, \dots, x_n)$ un point de R^n . On désigne par $W^{1,p}(\Omega)$, $1 < p < +\infty$ l'espace des (classes de) fonctions $u \in L^p(\Omega)$ telles que $D_i u = \frac{\partial u}{\partial x_i} \in L^p(\Omega)$; les dérivées étant prises au sens des distributions. $W_0^{1,p}(\Omega)$ désigne l'adhérence de $\mathcal{D}(\Omega)$ dans $W^{1,p}(\Omega)$.

On note par $W^{-1,q}(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$, le dual de $W_0^{1,p}(\Omega)$.

Posons $V = W_0^{1,p}(\Omega)$. On sait (cf. [9]) que, V coïncide avec

l'espace $\{u \mid u \in L^p(\Omega), D_i u \in L^p(\Omega), i = 1, \dots, n; u|_{\Gamma} = 0\}$

et $V' = W^{-1,q}(\Omega) = \{f \mid f = f_0 + \sum_{i=1}^n D_i f_i, f_0, f_i \in L^q(\Omega)\}$.

Muni de la norme $\|u\| = \left(\sum_{i=1}^n \|D_i u\|_{L^p(\Omega)}^p \right)^{1/p}$, V est un espace uniformément convexe.

Soit X un convexe de V contenant l'origine.

On pose : $A u = - \sum_{i=1}^n D_i (|D_i u|^{p-2} D_i u) - \lambda \Delta u \quad \lambda > 0 \quad p \geq 2$

et l'on cherche $u \in X$ tel que

$$(3.1) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

pour f donné dans V' .

Nous avons montré dans [14] chap. II, que le problème (3.1) admettait une solution unique. Le problème (3.1) se discrétise alors (cf. [14] chap. II) sous la forme :

$$(3.2) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

avec

$$(A_h u_h, v_h)_h = \sum_{i=1}^n \int_{\Omega} |\nabla_i u_h|^{p-2} \nabla_i u_h \nabla_i v_h dx + \lambda \sum_{i=1}^n \int_{\Omega} \nabla_i u_h \nabla_i v_h dx$$

$$\nabla_i u_h(x) = \frac{u_h(x + \frac{h_i}{2}) - u_h(x - \frac{h_i}{2})}{h_i}$$

$$f_h = \frac{1}{h_1 \dots h_n} \int_{\Omega} f \theta_h^M dx \quad \text{où } \theta_h \text{ est la fonction caractéristique}$$

d'un pavé de centre x et de côtés $h_1, \dots, h_n$ respectivement parallèles aux axes de coordonnées $0x_1 \dots 0x_n$.

Nous avons vu dans [14] (cf. théorème 1.4 chap. II) que $p_h^i u_h \rightarrow u_h$ dans $W_i^{1,p}(\Omega)$ fort pour $i=1, \dots, n$; u_h étant la solution de (3.2) dans $X_h \subset V_h$, espace de dimension fini associé au paramètre $h = (h_1, \dots, h_n)$ de discrétisation; et u étant la solution du problème (3.1).

On se propose maintenant de résoudre numériquement le problème (3.2) discrétisé en précisant chaque fois les convexes X et X_h .

III-1 Problème 1 : le convexe X est une boule

$$\text{Posons } X = \{u \mid u \in W_0^{1,p}(\Omega) ; \sum_{i=1}^n \|D_i u\|_{L^2(\Omega)}^2 \leq 1\}$$

On cherche $u \in X$ solution de

$$(3.1) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

pour f donné dans $W^{-1,q}(\Omega)$ avec

$$A u = - \sum_{i=1}^n D_i (|D_i u|^{p-2} D_i u) - \lambda \Delta u \quad \lambda > 0$$

Dans un espace $V_h(\Omega_h)$ de dimension finie (cf. [14]) (3.1) s'écrit sous la forme (3.2) avec

$$(3.3) \quad X_h = \{u_h \mid u_h \in V_h(\Omega_h) ; \sum_{i=1}^n \|\nabla_i u_h\|_{L^2(\Omega)}^2 \leq 1\}.$$

Nous allons maintenant résoudre par itérations le problème (3.2), (3.3).

Posons

$$S_h u_h = - \sum_{i=1}^n \nabla_i^2 u_h.$$

On munit $V_h(\Omega_h)$ du nouveau produit scalaire

$$[u_h, v_h] = (S_h u_h, v_h)_h = \sum_{i=1}^n (\nabla_i u_h, \nabla_i v_h)_{L^2(\Omega)}$$

et de la norme associée $[u_h] = \left(\sum_{i=1}^n \|\nabla_i u_h\|_{L^2}^2 \right)^{1/2}$

On définit alors

$$P_{X_h}^S(v_h) = \begin{cases} v_h & \text{si } [v_h] \leq 1 \\ \frac{v_h}{[v_h]} & \text{si } [v_h] > 1 \end{cases}$$

et l'on cherche une suite (u_h^n) approchant la solution $u_h \in X_h$ (X_h étant défini par (3.3)) de l'inéquation

$$(3.2) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h \quad \text{avec}$$

$$(3.4) \quad A_h u_h = - \sum_{i=1}^n \nabla_i (|\nabla_i u_h|^{p-2} \nabla_i u_h) - \lambda \sum_{i=1}^n \nabla_i^2 u_h.$$

Nous avons alors le

Théorème 3.1.

1 - La suite (u_h^n) définie par l'itération

$$(3.5) \quad u_h^{n+1} = P_{X_h}^{S_h} (u_h^n - \rho S_h^{-1} (A_h u_h^n - f_h))$$

où X_h et A_h sont donnés par (3.3) et (3.4); converge vers

u_h solution de (3.2) quand $n \longrightarrow \infty$ pour $u_h^0 = 0$,

$\rho = \lambda / C^2(N)$ et

$$C(N) = (p-1) \frac{N^{p-2}}{(h_1, \dots, h_n)^{\frac{p-2}{2}}} + \lambda ; \quad \text{la constante } N \text{ pouvant}$$

être prise égale à $\frac{2}{\lambda \sqrt{2}} \|f_h\|_{L^2(\Omega)}$

2 - De plus on a $P_h^i u_h \longrightarrow u$ dans $W_i^{1,p}(\Omega)$ fort $i = 0, 1, \dots, n$ quand $h \longrightarrow 0$; où u est la solution (2.1) avec

$$X = \{u \mid u \in W_0^{1,p}(\Omega); \sum_{i=1}^n \|D_i u\|_{L^2(\Omega)}^2 \leq 1\}.$$

Démonstration

On se place dans le cadre du théorème 1.1 chap. I..

1 - On vérifie facilement la relation

$$(3.6) \quad (A_h u_h - A_h v_h, u_h - v_h)_h \geq \lambda [u_h - v_h]^2 \quad u_h, v_h \in V_h$$

Calcul de C(N)

Supposons que $[u_h] \leq N$, $[v_h] \leq N$. On a alors :

$$[u_h]^2 = \sum_{i=1}^n \|\nabla_i u_h\|_{L^2}^2 = \sum_{i=1}^n h_i \dots h_n \sum_x |\nabla_i u_h(x)|^2$$

Donc

$$|\nabla_i u_h(x)| \leq \frac{[u_h]}{(h_1 \dots h_n)^{1/2}} = N' \quad \forall x \text{ et } \forall i.$$

Par ailleurs

$$\begin{aligned} |(A_h u_h - A_h v_h, w_h)_h| &\leq \left(\sum_{i=1}^n \|\nabla_i u_h\|_{L^2}^{p-2} \nabla_i u_h - \|\nabla_i v_h\|_{L^2}^{p-2} \nabla_i v_h \right)^{1/2} \left(\sum_{i=1}^n \|\nabla_i w_h\|_{L^2}^2 \right)^{1/2} \\ &\quad + \lambda \sum_{i=1}^n \|\nabla_i (u_h - v_h)\|_{L^2} \|\nabla_i w_h\|_{L^2} \end{aligned}$$

Ce qui donne

$$|(A_h u_h - A_h v_h, w_h)_h| \leq C(N) [u_h - v_h] [w_h]$$

En posant :

$$(3.7) \quad C(N) = (p-1) N^{p-2} + \lambda = (p-1) \frac{N^{p-2}}{(h_1 \dots h_n)^{\frac{p-2}{2}}} + \lambda$$

Calcul de N

On a successivement :

$$\sum_{i=1}^n \|\nabla_i u_h\|_{L^p}^p \leq (A_h u_h, u_h) \leq \|f_h\|_{L^q} \|u_h\|_{L^p}$$

$$\sum_{i=1}^n \|\nabla_i u_h\|_{L^p}^p \leq C_1^{1/p}(\Omega) \|f_h\|_{L^q} \left(\sum_{i=1}^n \|\nabla_i u_h\|_{L^p}^p \right)^{1/p}$$

où $C_1(\Omega)$ est une constante ne dépendant que de l'ouvert Ω .

Comme $L^p \subset L^2$ avec une injection continue pour $p \geq 2$; on a :

$$[u_h] = \left(\sum_{i=1}^n \|\nabla_i u_h\|_{L^2}^2 \right)^{1/2} \leq \left(\sum_{i=1}^n \|\nabla_i u_h\|_{L^p}^p \right)^{1/p} \leq C_1^{q/p}(\Omega) \|f_h\|_{L^q}^{1/p} = N$$

Donc $N = c^{te}$ indépendante de h .

En posant $\rho = \lambda / C^2(N)$ les itérations (3.5) sont convergentes d'après le théorème 1.1 chap. I.

2 - Cette assertion découle du théorème 1.4 chap. II de [14].

Remarque 3.1

1 - D'après le théorème 1.1 chap. I on peut aussi prendre

$$N = \frac{2}{\lambda \sqrt{\sigma}} \|f_h\|_{L^2} \quad \text{avec } \sigma \quad \text{tel que} \quad [u_h]^2 \geq \sigma \|u_h\|_{L^2}^2$$

2 - La formule (1.6) du chap. I donne les majorations de l'erreur suivantes :

$$C_1(\Omega) \|u_h^n - u_h\|_{L^2}^2 \leq \sum_{i=1}^n \|\nabla_i u_h^n - \nabla_i u_h\|_{L^2}^2 \leq \theta^n \|u_h\|_{L^2}^2$$

où $C_1(\Omega)$ est une constante ne dépendant que de l'ouvert Ω

et $\theta = 1 - \lambda^2 / C^2(N)$

3 - Si $n \sim \frac{-2 \operatorname{Log} \|u\| + 2 \operatorname{Log} \epsilon}{\operatorname{Log} \theta}$ on a alors

$$\|u_h^n - u_h\|_{L^2} < \epsilon / C_1(\Omega) \quad \text{et} \quad \sum_{n=1}^n \|\nabla_i u_h^n - \nabla_i u_h\|_{L^2}^2 < \epsilon$$

Dans la pratique ces majorations sont réalisées pour des valeurs de n bien inférieures .

Remarque 3.2

1 - Si $\Omega = (0,1)$ $h > 0$ on montre par un calcul facile que l'on peut prendre selon la convenance

$$N = \frac{2}{\lambda} \|f_h\|_{L^2}^2 \quad \text{ou} \quad N = \|f\|_{L^q}^{1/p} \quad \text{avec} \quad \frac{1}{p} + \frac{1}{q} = 1$$

$$\text{et} \quad C(N) = (p-1) \frac{N^{p-2}}{h^{\frac{p-2}{2}}} + \lambda \quad p \geq 2.$$

2 - Si $\Omega = (0,1) \times (0,1)$ et $h_1 = h_2 = h$ on peut prendre

$$N = \frac{\sqrt{2}}{\lambda} \|f_h\|_{L^2}^2 \quad \text{ou} \quad N = \frac{1}{2^{q/p-2}} \|f_h\|_{L^q}^{1/q}$$

$$\text{et } C(N) = (p-1) \frac{N^{p-2}}{h} + \lambda$$

3 - D'autres estimations de la constante N conduisant à un meilleur coefficient de contraction $\theta = 1 - \frac{\lambda^2}{C^2(N)}$ sont possibles (cf. chap. I, 1.2)

Résultats numériques : un exemple unidimensionnel

On cherche $u \in X \subset W_0^{1,3}(0,1)$ vérifiant

$$(3.8) \quad (Au, v-u) \geq (f, v-u) \quad \forall v \in X \text{ pour } f \text{ donné dans } L^{3/2}(0,1)$$

avec

$$Au = -\frac{d}{dx} \left(\left| \frac{du}{dx} \right| \frac{du}{dx} \right) - \lambda \frac{d^2 u}{dx^2}$$

$$X = \left\{ u \mid u \in W_0^{1,3}(0,1), \quad \left\| \frac{du}{dx} \right\|_{L^2} \leq 1 \right\}$$

Au problème (3.8) on associe le problème discrétisé

$$(3.9) \quad \begin{cases} (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h ; f_h = f \\ A_h u_h = -\nabla (|\nabla u_h| \nabla u_h) - \lambda \nabla^2 u_h \\ X_h = \left\{ u_h \in V_h \mid \|\nabla u_h\|_{L^2} \leq 1 \right\} \end{cases}$$

et l'itération

$$(3.10) \quad \begin{cases} u_h^{n+1} = P_{X_h}^{S_h} (u_h^n - \rho S_h^{-1} (A_h u_h^n - f_h)) ; \quad u_h^0 = 0 \\ \text{avec } S_h = -\nabla^2, \quad \rho = \frac{\lambda}{C^2(N)} \quad \text{et} \quad C(N) = \frac{4 \|f_h\|_{L^2}^2}{\lambda \sqrt{h}} + \lambda \end{cases}$$

Cherchons une solution exacte de la forme

$$u(x) = a(x - x^2) \quad x \in (0,1) \quad a > 0$$

Elle doit bien entendu vérifier l'équation

$$(3.11) \quad u = P_X^S [u - \rho \left(-\frac{du}{dx} \right)^{-1} (Au - f)] ; \quad S = \frac{d^2}{dx^2}$$

Posons alors $f = Au + 2$

ce qui donne $\left(-\frac{du}{dx}\right)^{-1} (Au - f) = -(x - x^2)$

Ce qui donne

$$P_X^S \{ u - \rho S^{-1} (Au - f) \} = P_X^S \{ (a + \rho)(x - x^2) \}$$

$$\text{or } [(a + \rho)(x - x^2)] = (a + \rho) \left\| \frac{d}{dx} (x - x^2) \right\|_{L^2} = \frac{a + \rho}{\sqrt{3}} > 1$$

si on pose $a = \sqrt{3}$ on a alors

$$P_X^S \{ u - \rho S^{-1} (Au - f) \} = \sqrt{3} (x - x^2) = u \quad \text{ce qui vérifie (2.4)}$$

Nous prenons donc $u = \sqrt{3} (x - x^2)$ comme solution exacte et $f = Au + 2$. Nous calculons alors par nos méthodes (u_h^n) ($h \rightarrow 0$, $n \rightarrow \infty$) que nous comparons à la solution $u(x) = \sqrt{3} (x - x^2)$. L'opérateur $S_h = -\nabla^2$ est inversé à l'aide de la méthode bien connue de Gauss-Seidel [16]; les itérations internes devant se poursuivre jusqu'à ce que

$$\left(\sum_i | \bar{u}^{m+1}(i) - \bar{u}^m(i) |^2 \right)^{1/2} \leq 10^{-4}.$$

Nous obtenons les résultats suivants pour $\lambda = 10$:

$$\|u_h^n - u\|_{L_2} = 2 \cdot 10^{-3} \quad \text{pour } h = 1/10$$

$$\|u_h^n - u\|_{L_2} = 5 \cdot 10^{-4} \quad \text{pour } h = 1/20$$

Nous constatons que les résultats sont un peu moins précis que dans le problème 2 pourtant plus compliqué. Cela tient à la manière dont on calcule la projection sur X_h (cf. remarque 3.4).

Pour ce problème à une dimension nous n'avons pas cherché à obtenir des "performances" numériques (cf. problème 2 en dimension $n = 2$).

Remarque 3.3

1) L'étude des majorations à priori de la solution exacte u , conduit à une meilleure estimation du coefficient de contraction θ correspondant aux itérations et par conséquent à une estimation du nombre d'itérations à effectuer pour avoir $[u_h^n - u] < \epsilon$ pour ϵ donné à l'avance. Si l'on a des renseignements sur la norme $[u]$, le nombre d'itérations pour avoir la solution diminue grandement. Mais le problème peut toujours être résolu rien qu'avec des estimations à priori de $[u]$ obtenues à partir de l'opérateur A .

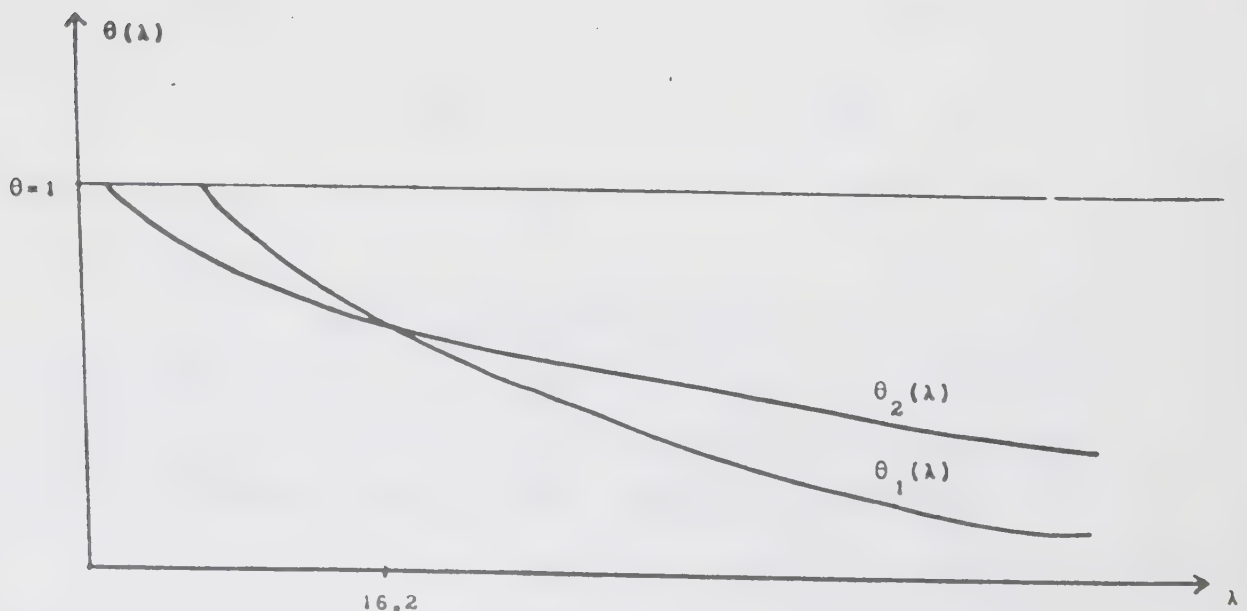
Nous donnons à titre d'exemple l'évolution du coefficient de contraction pour les itérations (3.10) en fonction de λ pour h fixé ($h = \frac{1}{10}$) selon que l'on estime $[u]$ par

$$N_1 = \frac{2}{\lambda} \|f\|_{L^2} \quad \text{ce qui conduit à} \quad \theta_1 = 1 - \frac{\lambda^2}{\left(\frac{4}{\lambda\sqrt{h}} \|f\|_{L^2} + \lambda\right)^2}$$

ou par

$$N_2 = \|f\|_{L^{3/2}}^{1/2} \quad \text{ce qui conduit à} \quad \theta_2 = 1 - \frac{\lambda^2}{\left(\frac{2}{\sqrt{h}} \|f\|_{L^{3/2}}^{1/2} + \lambda\right)^2}$$

Nous voyons que la majoration $[u] < N_2$ est plus avantageuse pour $\lambda < 16$. Pour λ assez grand les deux majorations à priori sont équivalentes.



2) Pour d'autres estimations de N cf remarque 3.5.

Résultats numériques : Exemple bidimensionnel.

Soit $\Omega = (0,1) \times (0,1)$ on cherche

$$(3.12) \quad u \in X = \left\{ u \mid u \in W_0^{1,3}(\Omega) ; \left(\sum_{i=1}^2 \|D_i u\|_{L^2}^2 \right)^{1/2} \leq 1 \right\}$$

vérifiant l'équation

$$(3.13) \quad (Au, v-u) \geq (f, v-u) \quad \forall v \in X \text{ pour } f \text{ donné dans } L^{3/2}(\Omega)$$

$$\text{avec } Au = - \sum_{i=1}^2 D_i (|D_i u| D_i u) - \lambda \Delta u \quad \lambda > 0$$

Au problème (3.13) on associe le problème discrétisé

$$(3.14) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall u_h \in X_h$$

avec

$$(3.15) \quad X_h = \left\{ u_h \in V_h \mid \left(\sum_{i=1}^2 \|\nabla_i u_h\|_{L_2}^2 \right)^{1/2} \leq 1 \right\}$$

$$A_h u_h = - \sum_{i=1}^2 \nabla_i (|\nabla_i u_h| \nabla_i u_h) - \lambda \sum_{i=1}^2 \nabla_i^2 u_h$$

et l'itération

$$(3.16) \quad u_h^{n+1} = P_{X_h}^{S_h} (u_h^n - \rho S_h^{-1} (A_h u_h^n - f_h))$$

$$\text{où } u_h^0 = 0, \quad S_h = - \sum_{i=1}^2 \nabla_i^2, \quad \rho = \lambda / C^2(N) \text{ et } C(N) = \frac{2\sqrt{2}}{\lambda h} \|f_h\|_{L_2} + \lambda$$

Les mêmes considérations qu'en dimension $n = 1$ (problème 1) conduisent à prendre comme solution exacte

$$(3.17) \quad u = \sqrt{45} (x - x^2) (y - y^2) \quad (x, y) \in (0,1) \times (0,1)$$

Posons alors

$$(3.18) \quad f_h = f(x, y) = Au + 2(x - x^2) + 2(y - y^2)$$

Nous calculons par (3.16) la suite (u_h^n) que nous comparons à u donné par (3.17).

L'opérateur $S_h = - \sum_{i=1}^2 v_i^2$ est inversé par la méthode de relaxation successive cf. [16].

Remarque 3.4

On a choisi la solution exacte telle que $[u] = 1$. Quand on calcule $[u]$ par la machine on trouve $[u] = a$ voisin de 1 suivant les valeurs du pas h . Aussi dans les itérations (3.16) on définit la projection de la manière suivante ;

Posons : $v_h^n = u_h^n - \rho S_h^{-1} (A_h u_h^n - f_h)$

Si $[v_h^n] \leq a$ alors $u_h^{n+1} = v_h^n$

Si $[v_h^n] > a$ alors $u_h^{n+1} = v_h^n \frac{a}{[v_h^n]}$

ce qui élimine l'erreur sur le calcul de la norme $[\]$.

Nous obtenons alors les résultats suivants :

1) Pour $\lambda = 10$, $h = 1/10$ on a

$$\|u_h^n - u\|_{L^2} = 4 \cdot 10^{-5} \quad \text{avec } n = 70 \text{ itérations}$$

2) Pour $\lambda = 10$, $h = 1/15$ on a

$$\|u_h^n - u\|_{L^2} = 2 \cdot 10^{-6} \quad \text{avec } n = 80 \text{ itérations.}$$

Remarque 3.5

Nous pouvons aussi améliorer l'estimation du paramètre ρ de la manière suivante :

on a vu que

$$\rho = \lambda / C(N) \quad \text{avec} \quad C(N) = 2 \frac{N}{h} + \lambda$$

et N est une constante telle que $[u] \leq N$.

Nous pouvons aussi écrire :

$$(3.19) \quad \rho = \lambda / C(N') \quad \text{avec} \quad C(N') = 2 N' + \lambda \quad (\text{cf. formule (3.7)})$$

avec $\sup |\nabla_i u_h| \leq N'$

nous pouvons alors poser :

$$(3.20) \quad \rho = \rho_n = \lambda / C^2(N'_n) \quad C(N'_n) = 2 N'_n + \lambda$$

en prenant $N'_n = \sup |\nabla_i u_h^{n-1}|$

Ceci améliore beaucoup les résultats puisqu'on a :

pour $\lambda = 10$, $h = 1/10$ $\|u_h^n - u\|_{L_2} = 8,4 \cdot 10^{-6}$ avec $n = 18$

itérations et après, les résultats sont stationnaires.

L'erreur relative est de l'ordre de 10^{-5} . 44 s suffisent en C.D.C. 3.600 pour exécuter 50 itérations (compilation + exécution).

III-2 Problème 2 : régularisation elliptique.

On veut résoudre dans $W_0^{1,3}(\Omega)$ l'inégalité :

$$(3.21) \quad (Bu, v - u) \geq (f, v - u) \quad \forall v \in X$$

avec
$$Bu = - \sum_{i=1}^n D_i (|D_i u| D_i u)$$

où X est donné par (3.12).

Le théorème 4.1 chap. I de [4], montre qu'il suffirait de résoudre (3.13) pour λ assez petit.

On a alors $u_{\lambda} \xrightarrow[\lambda \rightarrow 0]{} u$ (solution de (3.21)) dans $W_0^{1,3}(\Omega)$ fort.

Nous utilisons alors les itérations (3.16) en posant

$$(3.22) \quad \rho = \rho_n = \frac{k \lambda}{2 N'_n + k \lambda} \quad \text{avec} \quad N'_n = \sup |\nabla_i u_h^{n-1}|$$

où k est une constante à choisir.

(Ce paramètre ρ est en fait une variante de (3.20)).

Nous obtenons alors les résultats suivants :

Pour $h = 1/10$ (domaine à 100 points intérieurs)

Il suffit de prendre $\lambda = 1/100$ ($k = 1000$) pour avoir au bout de $n = 900$ itérations

$$\|u_{h,\lambda}^n - u\|_{L_2} = 1,10 \cdot 10^{-5}$$

où u est la solution exacte de (3.21), (3.12) .

6 mn 25 s suffisent pour le calcul de 1000 itérations en C.D.C. 3 600 (exécution + compilation).

Nous avons donc en fait les convergences suivantes :

- 1) $u_{h,\lambda}^n$ solution de (3.16) converge vers $u_{h,\lambda}$ solution de (3.14)
- 2) $P_h u_{h,\lambda} \xrightarrow{\lambda \rightarrow 0} u_\lambda$ où u_λ est la solution de (3.13)
- 3) $u_\lambda \xrightarrow{\lambda \rightarrow 0} u$ où u est la solution de (3.21).

Remarque 3.6

Cet exemple montre qu'on peut encore avec un choix optimum du paramètre $\rho = \rho_n$ se passer encore des méthodes à double itérations du chap. I paragraphe I-3.

III-3 Problème 3 : le convexe X est cône.

Soit E un compact de Ω . Posons le même problème que dans l'exemple III-1 en prenant cette fois comme convexe :

$$X = \{u \mid u \in W_0^{1,p}(\Omega), \quad u \geq 0 \quad \text{p.p sur } E\}$$

Un problème analogue pour l'opérateur du 2ème ordre est posé dans [10]
On montre facilement que

$$(3.23) \quad X_h = \{u_h \in V_h \mid u_h(x) \geq 0 \quad \text{p.p sur } E\}$$

Il s'agit de trouver $u_h \in X_h$ solution de

$$(3.24) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

$$\text{avec} \quad A_h u_h = - \sum_{i=1}^n \nabla_i (|\nabla_i u_h|^{p-2} \nabla_i u_h) - \lambda \sum_{i=1}^n \nabla_i^2 u_h \quad \lambda > 0.$$

Nous désignons par R_h le réseau des points M de la forme

$$M = (m_1 h_1, \dots, m_n h_n) \quad ; \quad m_i \in \mathbb{Z}, \quad i = 1, \dots, n; \quad h_i > 0$$

Nous posons

$$\omega_{h,q}^M = \prod_{i=1}^n \left[\left(m_i - \frac{q_i + 1}{2} \right) h_i, \left(m_i + \frac{q_i + 1}{2} \right) h_i \right]$$

avec q_i entier positif non nul.

$$\rho_h^M = \bigcup_{|q| \leq 1} \omega_{h,q}^M \quad \text{et} \quad \Omega_h = \{M \mid M \in \mathcal{R}_h, \rho_h^M \subset \Omega\}$$

$$V_h = \{u_h \mid u_h = \sum_{M \in \Omega_h} \xi_M \mathcal{W}_M, \xi_M \in \mathbb{R}\} \quad \text{où } \mathcal{W}_M \text{ est la}$$

fonction caractéristique du pavé de centre M de côtés $h_1, \dots, h_n$ respectivement parallèles aux axes de coordonnées $ox_1, \dots, ox_n$.

Pour tout $u_h \in V_h$ on peut écrire :

$$u_h = \sum_{M \in \Omega_h^1} \xi_M \mathcal{W}_M + \sum_{M \in \Omega_h^2} \eta_M \mathcal{W}_M$$

$$\text{où} \quad \Omega_h^1 = \{M \mid M \in \mathcal{R}_h, \rho_h^M \cap E \neq \emptyset\}$$

$$\Omega_h^2 = \Omega_h - \Omega_h^1$$

$$\text{Soit alors } v_h = \sum_{M \in \Omega_h^1} \xi_M^* \mathcal{W}_M + \sum_{M \in \Omega_h^2} \eta_M^* \mathcal{W}_M \quad \text{alors}$$

$$v_h \in X_h \iff \xi_M^* \geq 0 \quad \text{p.p sur } \Omega_h^1.$$

Proposition 3.1

Si $u_h \in V_h$ alors

$$P_{X_h}(u_h) = \begin{cases} \eta_M & \text{si } M \in \Omega_h^2 \\ \xi_M & \text{si } M \in \Omega_h^1 \\ 0 & \text{si } M \in \Omega_h^1 \end{cases} \quad \begin{matrix} \text{et } \xi_M \geq 0 \\ \text{et } \xi_M < 0 \end{matrix}$$

Démonstration

Nous avons en effet à minimiser l'expression

$$\|u_h - v_h\|^2 = \sum_{H \in \Omega_h} |\xi_H - \xi'_H|^2 + \sum_{H \in \Omega_h} |\eta_H - \eta'_H|^2$$

ce qui donne $\eta'_H = \eta_H$

$$\text{et } \xi'_H = \begin{cases} \xi_H & \text{si } \xi_H \geq 0 \\ 0 & \text{si } \xi_H < 0 \end{cases}$$

D'où le

Théorème 3.2

1 - La suite (u_h^n) définie par les itérations :

$$(3.25) \quad u_h^{n+1} = P_{X_h}^{S_i} (u_h^n - \rho_i S_i^{-1} (A_h u_h^n - f_h))$$

$$\text{où } X_h = \{u_h \in V_h \mid u_h(x) \geq 0 \text{ p.p sur } E\}$$

converge vers u_h solution de (3.24)

pour u_h^0 et S_i convenables (cf. démonstration du théorème).

2 - De plus on a $P_h^i u_h \longrightarrow u$ dans $W_0^{1,p}(\Omega)$ fort quand $h \longrightarrow 0$ où u est la solution de (3.1) avec

$$X = \{u \mid u \in W_0^{1,p}(\Omega), \quad u \geq 0 \text{ p.p sur } E\}$$

Démonstration

1°)

$$\text{a) Posons } S_i u_h = S_1 u_h = - \sum_{i=1}^n \frac{2}{\nabla_i} u_h$$

$$\rho_i = \rho_1 = \lambda / C_1^2(N) \quad \text{avec } C_1(N) = (p-1) \frac{N^{p-2}}{(h_1 \dots h_n)^{\frac{p-2}{2}}} + \lambda$$

$$\text{et } N = \frac{2}{\lambda \sqrt{\sigma}} \|f_h\|_{L_2} \quad \text{si } u_h^0 = 0$$

si $f_h^1 = u_h^n - \rho_1 S_1^{-1} (A_h u_h^n - f_h) \in X$ alors on a

$$u_h^{n+1} = f_h^1 \quad \text{et les itérations (3.25) convergent.}$$

b) si $f_n^1 = u_h^n - \rho_1 S_1^{-1} (A_h u_h^n - f_h) \notin X_h$ alors

on pose

$$S_1 u_h = S_2 u_h = u_h \quad \text{et} \quad [u_h] = \|u_h\|_{L_2}$$

on démontre facilement que

$$(A_h u_h - A_h v_h, u_h - v_h)_h \geq \lambda \sigma \|u_h - v_h\|^2$$

car

$$[u_h]^2 = \sum_{i=1}^n \|\nabla_i u_h\|_{L_2}^2 \geq \sigma \|u_h\|_{L_2}^2$$

De même si

$$\|u_h\| \leq N_0 \quad \text{et} \quad \|v_h\| \leq N_0 \quad \text{on a :}$$

$$\|u_h\|^2 = h_1 \dots h_n \sum_{M \in \Omega_h} |u_h(M)|^2$$

ce qui entraîne

$$\|u_h\|^2 \geq h_1 \dots h_n \sup_M |u_h(M)|^2$$

De même nous avons

$$|\nabla_i u_h(x)| \leq \frac{2}{h_i} \sup_y u_h(y)$$

D'où

$$|\nabla_i u_h(x)| \leq \frac{2}{h_i (h_1 \dots h_n)^{1/2}} \|u\| \leq \frac{2N_0}{h_i (h_1 \dots h_n)^{1/2}} = N'$$

Par ailleurs

$$\begin{aligned} (A_h u_h - A_h v_h, w_h)_h &\leq \left(\sum_{i=1}^n \|\nabla_i u_h\|^{p-2} \|\nabla_i u_h - \nabla_i v_h\|^{p-2} \right)^{1/2} \left(\sum_{i=1}^n \|\nabla_i w_h\|^2 \right)^{1/2} \\ &\quad + \lambda \left(\sum_{i=1}^n \|\nabla_i (u_h - v_h)\|^2 \right)^{1/2} \left(\sum_{i=1}^n \|\nabla_i w_h\|^2 \right)^{1/2} \end{aligned}$$

or pour $|\alpha| \leq N'$, $|\beta| \leq N'$ α, β réels on a :

$$||\alpha|^{p-2} \alpha - |\beta|^{p-2} \beta| \leq (p-1) N'^{p-2} |\alpha - \beta|$$

ce qui donne

$$\sum_{i=1}^n \|\nabla_i u_h\|^{p-2} \|\nabla_i u_h - \nabla_i v_h\|^{p-2} \leq (p-1) \sum_{i=1}^n N_i'^{p-2} \|\nabla_i (u_h - v_h)\|$$

Donc

$$(A_h u_h - A_h v_h, w_h)_h \leq C(N') \|u_h - v_h\| \|w_h\| \quad \text{avec}$$

$$C_2(N') = (p-1) \left(\sum_{i=1}^n \frac{N_i'^2 (p-2)}{h_i^2} \right)^{1/2} + \left(\sum_{i=1}^n \frac{4}{h_i^2} \right)^{1/2} + \lambda \left(\sum_{i=1}^n \frac{4}{h_i^2} \right)$$

$$N_i' = \frac{2 N_0}{h_i (h_1 \dots h_n)^{1/2}}$$

D'où

$$(A_h u_h - A_h v_h, w_h)_h \leq C_2(N_0) \|u_h - v_h\| \|w_h\| \quad \text{avec}$$

$$C_2(N) = (p-1) N_0^{p-2} \frac{2^p}{(h_1 \dots h_n)^{\frac{p-2}{2}}} \left(\sum_{i=1}^n \frac{1}{h_i^{2p-2}} \right)^{1/2} + \lambda \left(\sum_{i=1}^n \frac{4}{h_i^2} \right)$$

Alors les itérations (3.25) convergent d'après le théorème 1.1 pour

$$\rho_1 = \rho_2 = \lambda^0 / C_2^2(N) \quad N = \frac{2}{\lambda \sqrt{2}} \|f_h\|_{L_2} \quad \text{si } u_h^0 = 0$$

et d'après la proposition 3.1, on sait projeter

$$f_h^2 = u_h^n - \rho_2 (A_h u_h^n - f_h) \quad \text{suivant la norme } \|u_h\|_h = \|u_h\|_{L_2}$$

D'où la détermination de u_{n+1} à l'aide de (3.25).

De plus, on voit que les itérations (3.25) convergent plus rapidement avec $S_1 = S_1$ et $\rho_1 = \rho_1$ qu'avec le choix $S_1 = S_2 = Id$ et $\rho_1 = \rho_2$.

Remarque 3.7

On peut approcher le problème

$$(3.26) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

$$\text{où } X_h = \{ u_h \in V_h \mid u_h \geq 0 \text{ p.p. sur } E \}$$

par le problème

$$(3.27) \quad A_h u_h + \frac{1}{\epsilon} F'_h u_h = f_h$$

où ϵ est un paramètre positif destiné à tendre vers 0 ; et F'_h est la dérivée de Gateaux de la fonctionnelle

$$F_h(v_h) = \int_{\Omega} v_h^2 H_h dx \quad \text{avec} \quad H_h = \begin{cases} 0 & \text{si } v_h \geq 0 \\ 1 & \text{si } v_h < 0 \end{cases}$$

On sait alors (cf. théorème 6.1 de [14]) que $u_{h,\epsilon}$ solution de (3.27) converge fortement vers u_h solution de (3.26) quand $\epsilon \longrightarrow 0$.

Remarque 3.8

Soit λ_p une suite de réels tels que $\lambda_p \longrightarrow 0$ quand $p \longrightarrow \infty$ considérons l'inégalité

$$(3.28) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

$$\text{où} \quad A_h u_h = - \sum_{i=1}^n \nabla_i (|\nabla_i u_h|^{p-2} \nabla_i u_h)$$

Alors la suite $(u_{h,p}^{n+1})$ définie par les itérations :

$$(3.29) \quad u_{h,p}^{n+1} = P_{X_h}^{S_i} (u_{h,p}^n - \rho_i S_i^{-1} (A_h u_{h,p}^n - \lambda_p \sum_{i=1}^n \nabla_i^2 u_{h,p}^n - f_h))$$

converge dans V_h quand $n \longrightarrow \infty$ et $p \longrightarrow \infty$ vers u_h solution de (3.28) pour $u_{h,p}^0 = 0$ ρ_i et S_i étant choisis comme dans le théorème 3.2.

(Cf. théorème 3.1 au chapitre I.)

1ère expérience numérique.

Soit $\Omega = (0,1) \times (0,1)$ on cherche

$$(3.30) \quad u \in X = \left\{ u \in W_{\theta}^{1,3}(\Omega) \mid u \geq 0 \text{ p.p dans } \Omega \right\}$$

vérifiant l'inégalité (3.13).

On écrit alors les itérations

$$(3.31) \quad u_h^{n+1} = P_{X_h}^{S_i} (u_h^n - \rho_i S_i^{-1} (A_h u_h^n - f_h)) \quad ;$$

avec

$$A_h u_h = - \sum_{i=1}^2 \nabla_i (|\nabla_i u_h| \nabla_i u_h) - \lambda \sum_{i=1}^4 \nabla_i^2 u_h$$

$$X_h = \left\{ u_h \in V_h \mid u_h \geq 0 \right\}$$

1) Si $f_n = u_h^n - \rho_2 S_2^{-1} (A u_h^n - f_h) \in X_h$ où

$$S_2 = - \sum_{i=1}^2 \nabla_i^2 \quad ; \quad \rho_2 = \lambda / C_2^2(N) \quad ; \quad C_2(N) = \frac{2}{h} N + \lambda$$

$$\text{et } N = \frac{\sqrt{2}}{\lambda} \|f_h\|_{L_2}$$

alors on calcule $u_h^{n+1} = P_{X_h}^{S_2}(f_n) = f_n$

2) Si $f_n \notin X_h$ alors on fait $S_i = S_1 = Id$ et

$$\rho_i = \rho_1 = 2\lambda / C_1^2(N) \quad \text{avec} \quad C_1(N) = \frac{32\sqrt{2} N}{h^6} + \frac{8\lambda}{h^2} \quad N = \frac{\sqrt{2}}{2} \|f\|_{L_2}$$

et u_h^{n+1} est donné par la formule

$$u_h^{n+1} = P_{X_h} (u_h^n - \rho_1 (A_h u_h^n - f_h))$$

Pour l'expérience on se donne comme solution exacte

$$(3.32) \quad u = (x - x^2) (y - y^2)$$

et on pose

$$f_h = f = A u$$

On vérifie alors que l'on a bien

$$u = P_X^{S_i} (u - \rho_i S_i^{-1} (Au - f))$$

Nous calculons (u_h^n) (quand $n \rightarrow \infty$ et $h \rightarrow 0$) à l'aide des itérations (3.31) que nous comparons à u donné par (3.32).

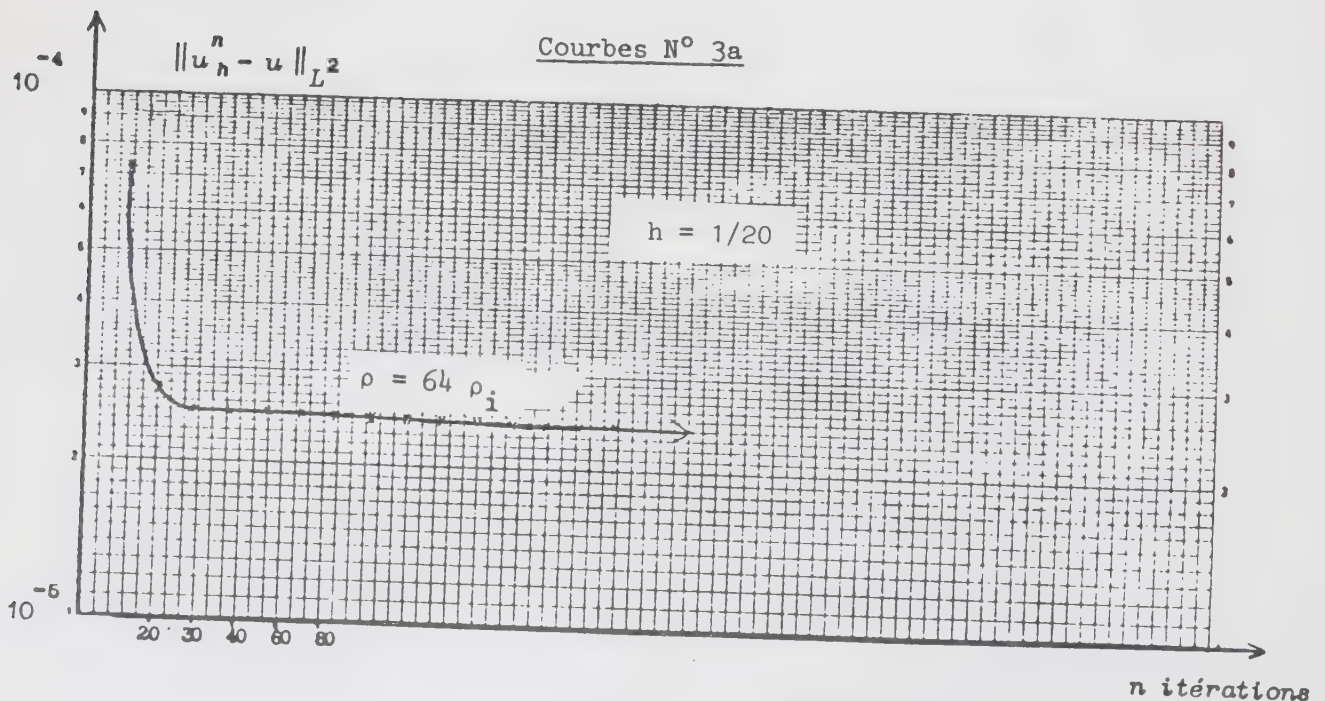
Pour être sûr que nous avons effectivement résolu l'inéquation (3.13) et non l'équation correspondante (cas où $X = V$) on vérifie ensuite qu'avec f quelconque on obtient une solution u_h^n qui réalise les contraintes de X .

Nous avons posé $h = 1/20$; ce qui donne un domaine Ω de 400 points intérieurs. Pour $\rho = \rho_i$ la convergence est assez lente. Pour $\rho = 8\rho_i$ (surestimé) la convergence est nettement plus rapide.

Enfin pour $\rho = 64\rho_i$ (cf. courbe n° 3a) on obtient au bout de $n = 20$ itérations $\|u_h^n - u\|_{L_2} = 2,5 \cdot 10^{-5}$, ensuite les résultats sont stationnaires.

L'exécution de 20 itérations en C.D.C. 3.600 dure 27,1s.

En posant $N_{1,n} = \|u_{n-1}\|$ et $N_{2,n} = [u_{n-1}]$ on est conduit à des paramètres $\rho_{i,1}^n$ et $\rho_{i,2}^n$ variables avec les itérations.



2ème expérience numérique.

Pour $\Omega = (0,1) \times (0,1)$ on cherche u vérifiant (3.30) et solution de l'inégalité (3.13) avec $\lambda = 1$.

On pose maintenant les itérations :

$$(3.33) \quad u_h^{n+1} = P_X^{S_i} (u_h^n - \rho_i^n S_i^{-1} (A_h u_h^n - f_h))$$

$$\text{avec} \quad A_h u_h = - \sum_{i=1}^2 \nabla_i (|\nabla_i u|^{p-2} \nabla_i u) - \lambda \sum_{i=1}^2 \nabla_i^2 u_h$$

$$X_h = \{u_h \in V_h \mid u_h \geq 0\}$$

et où le paramètre ρ_i^n varie avec les itérations.

Un calcul de variations montre que l'on a :

$$\pi^2 \int_{\Omega} |v|^2 dx \leq \int_{\Omega} \left| \frac{\partial v}{\partial x} \right|^2 dx \quad \text{pour} \quad v \in H_0^1(\Omega)$$

ce qui donne

$$\sigma \|v\|_{L^2}^2 \leq [v]^2 = \sum_{i=1}^2 \|D_i u\|_{L^2}^2 \quad \text{avec} \quad \sigma = 2 \pi^2$$

Nous procédons maintenant de la manière suivante

1) On pose dans (3.33)

$$S_i = S_2 = - \sum_{i=1}^2 \nabla_i^2 ; \quad \rho_2^n = \lambda / C_2^2(N_n) \quad \text{avec} \quad C_2(N_n) = \frac{2N}{h} + \lambda, \quad N_n = [u_h^n]$$

$$\text{si} \quad f_h = u_h^n - \rho_2^n S_2^{-1} (A u_h^n - f_h) \in X_h$$

$$\text{alors} \quad u_h^{n+1} = f_h$$

2) Si $f_h \in X_h$ alors on pose dans (3.33)

$$S_i = S_1 = \text{Id}, \quad \rho_1^n = \lambda \sigma / C_1^2(N_n) \quad \text{avec} \quad C_1(N_n) = \frac{64N}{h^4} + \frac{8\lambda}{h^2}, \quad N_n = \|u_h\|$$

et u_h^{n+1} est donné par la formule

$$u_h^{n+1} = P_{X_h}(u_h^n - \rho_1^n (A_h u_h^n - f_h)) .$$

Pour l'expérience numérique on se donne une solution exacte du problème (3.13), (3.30) par la formule :

$$(3.34) \quad u(x,y) = \frac{1}{2} (x^3 - x^2) (y^3 - y^2)$$

et on pose dans (3.33)

$$\begin{aligned} f_h = f = A u = & x(y^3 - y)^2 (3x-1) |3x-2| \\ & + y(x^3 - x)^2 (3y-1) |3y-2| \\ & - \lambda [(3x-1)(y^3 - y) + (3y-1)(x^3 - x)] \end{aligned}$$

Nous calculons alors u_h^n à l'aide des itérations (3.33) pour h suffisamment petit que nous comparons à u donné par (3.34).

Pour $h = 10^{-1}$ nous constatons qu'à partir de la 8ème itérations les paramètres ρ_1^n et ρ_2^n se stabilisent à leur valeurs optimales :

$$\rho_1^n = 0,798 \quad \text{et} \quad \rho_2^n = 0,249 \cdot 10^{-4}$$

Le tableau n° 3b donne l'évolution de l'erreur $\|u_n - u\| / \|u\|$ dans $L^2(\Omega)$ en fonction de $M = 1/h$ lorsque ρ_1^n varie à partir de la 2ème itération et ρ_2^n varie à partir de la 50ème (pour $n < 50$ $\rho_2^n = \rho_2 =$ valeur théorique).

Lorsqu'on augmente la dimension $M = 1/h$ le temps de calcul devient rapidement élevé car il est directement lié au nombre d'itérations nécessaires pour inverser le Laplacien $S_2 = -\Delta$ dans les itérations (3.33).

$$\text{ainsi pour } M = 20 \quad \frac{\|u - u_n\|}{\|u\|} = 1,15 \cdot 10^{-4}$$

$$M = 26 \quad \frac{\|u - u_n\|}{\|u\|} = 7,11 \cdot 10^{-5}$$

à la 40ème itération.

Alors que le nombre total d'itérations dans les diverses inversion du Laplacien est

pour	$M = 20$	$N = 1793$
pour	$M = 26$	$N = 2823$

Le tableau n° 3c donne l'évolution de l'erreur de Cauchy

$$\frac{\|u_n - u_{n-1}\|}{\|u\|} \quad \text{dans } L^2 \quad \text{en fonction de } M = 1/h \quad \text{pour } \epsilon = 10^{-6}$$

(précision dans l'inversion du Laplacien) et pour un nombre donné d'itérations.

3ème expérience numérique. Méthode de pénalisation.

On résoud l'équation (3.27) , à l'aide des itérations :

$$(3.35) \quad u_h^{n+1} = u_h^n - \rho_n S_h^n (C_h^\epsilon u_h^n - f_h)$$

$$\text{où } C_h^\epsilon u_h = - \sum_{i=1}^2 \nabla_i (|\nabla_i u_h| \nabla_i u_h) - \lambda \sum_{i=1}^2 \nabla_i^2 u_h + \frac{1}{\epsilon} F' u_h \quad \text{avec } \lambda = 2 .$$

$$S_h = - \sum_{i=1}^n \nabla_i^2$$

pour $n = 0; 1, \dots, 10$ on pose

$$\rho_1 = \rho_0 = \frac{2}{\lambda \sqrt{\sigma}} \|f_h - C_h^\epsilon(0)\|_h^* \quad \lambda = 2, \quad \sigma = 2 .$$

pour $n > 10$ on pose

$$\rho_n = \lambda / C^2(N_n) \quad \text{avec} \quad C(N_n) = \frac{2N}{h} + \lambda + \frac{1}{2\epsilon}$$

$$N_{n+1} = \frac{1}{\lambda \sqrt{\sigma}} \|f - C_h^\epsilon u_h^n\|^* + \sqrt{\sigma} \|u_h^n\|$$

Plus précisément on procède de la manière suivante :

1) Posons $h_1 = h_2 = h^0 = 1/10$; on calcule alors

$$u_{h^0, \epsilon_i}^n \quad i = 0, 1, \dots \quad \text{pour } \epsilon_0 = 2 \quad \text{on pose} \quad u_{h^0, \epsilon_0}^0 = 0$$

Nous avons alors $\lim_{n \rightarrow \infty} u_{h^0, \epsilon_0}^n = u_{h^0, \epsilon_0}$

pour $\epsilon_1 = 0,5$ on pose $u_{h^0, \epsilon_1}^0 = u_{h^0, \epsilon_0}^0$

Nous avons alors $\lim_{n \rightarrow \infty} u_{h^0, \epsilon_1}^n = u_{h^0, \epsilon_1}^0$

on refait les calculs pour $\epsilon_2 = 0,125$

$$\epsilon_3 = 0,313 \cdot 10^{-1} ; \quad \epsilon_4 = 0,781 \cdot 10^{-2} ; \quad \epsilon_5 = 0,195 \cdot 10^{-2} ;$$

$$\epsilon_6 = 0,488 \cdot 10^{-3} ; \quad \epsilon_7 = 0,305 \cdot 10^{-4} ; \quad \text{et} \quad \epsilon_8 = 0,763 \cdot 10^{-5}$$

ce qui donne alors

$$\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} u_{h^0, \epsilon}^n = u_{h^0}^0$$

2) On pose alors $h^1 = h^0/2 = 1/20$

On initialise $u_{h^1, \epsilon_0}^0 = \tilde{u}_{h^0}^0 =$ interpolation linéaire de $u_{h^0}^0$; et

on effectue les itérations (3.35) pour $\epsilon_1, \epsilon_2, \dots, \epsilon_8$.

3) Si nous posons dans (3.27) $f = A u$

$$\text{où} \quad u(x, y) = P_X [2(x - x^2) (y - y^2) (\frac{1}{4} - y) (\frac{1}{4} - x)]$$

on montre facilement que

$$\lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} u_h^n = u$$

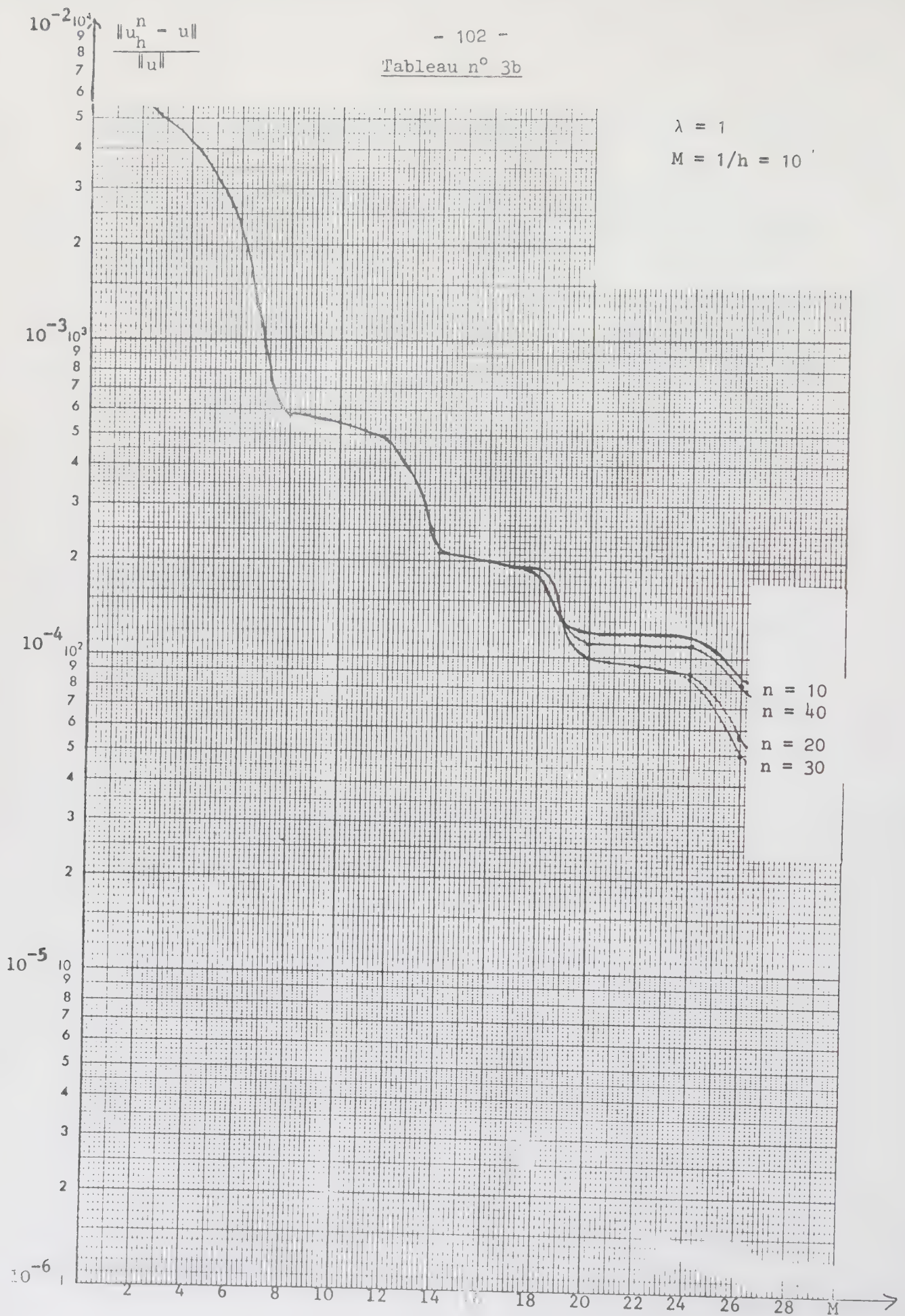
Les expériences numériques effectuées en double précision donnent les résultats suivants :

Pour $h^0 = 1/10$ nous obtenons le tableau suivant :

paramètre ϵ de pénalisation	erreur $\ u_h^\epsilon - u\ _{L_2}$
$\epsilon_1 = 0.500$	$0.6589 \cdot 10^{-19}$
$\epsilon_2 = 0.125$	$0.6065 \cdot 10^{-19}$
$\epsilon_3 = 0.313$	$0.6005 \cdot 10^{-19}$
.....	
$\epsilon_5 = 0.763 \cdot 10^{-5}$	$0.3593 \cdot 10^{-19}$

Pour $h^1 = 1/20$ nous obtenons $\|u_h^n - u\|_{L_2} = 0,86 \cdot 10^{-19}$ pour les mêmes valeurs du paramètre ϵ .

Tableau n° 3b



10^{-4}

Erreur de Cauchy

- 103 -

Tableau n° 3c

$\lambda = 1$

$\epsilon = 10^{-6}$

$M = 1/h$

10^{-5}

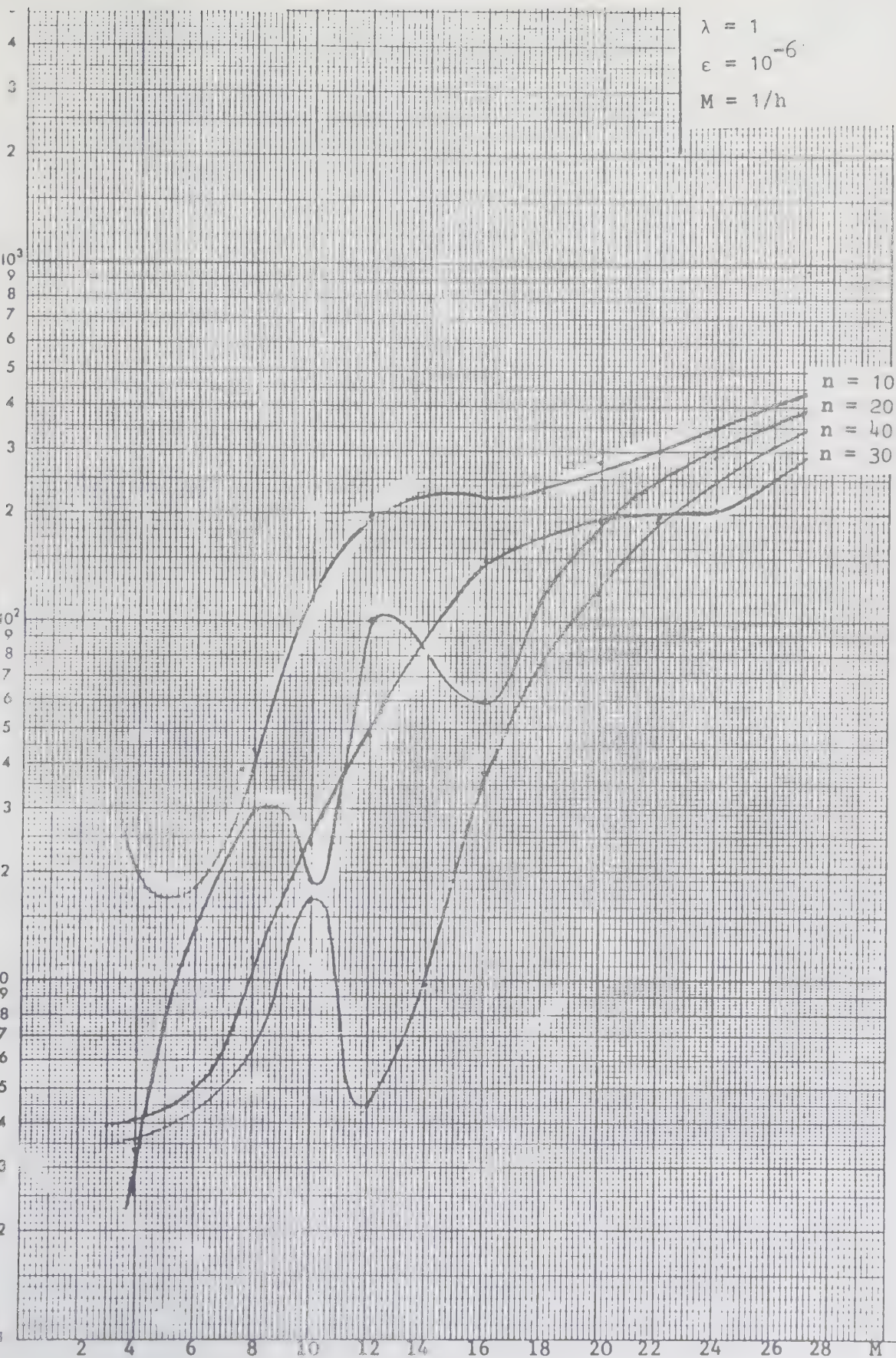
10^{-6}

10^{-7}

10^{-8}

$n = 10$
 $n = 20$
 $n = 40$
 $n = 30$

2 4 6 8 10 12 14 16 18 20 22 24 26 28 M →



Remarque 3.5

On cherche à résoudre le problème (3.13) dans le cas où le convexe X est de la forme

$$(3.36) \quad X = \left\{ u \in W_0^{1,3}(\Omega) \mid |u| \leq 1 \text{ p.p. dans } \Omega \right\}$$

On applique les itérations (3.25) dans lesquels X_h signifie :

$$X_h = \left\{ u_h \in V_h \mid |u_h| \leq 1 \text{ p.p.} \right\}$$

Dans une première étape on se donne une solution exacte

$$(3.37) \quad u = (x - x^2)(y - y^2)$$

et on pose :

$$f_h = f = Au.$$

On calcule par (3.25) u_h^n que l'on compare à u donné par (3.37)

1) Pour $h = 1/10$ et $p = p_i$

Nous obtenons à partir de $n = 50$ itérations

$$\|u_h^n - u\|_{L^2}^2 = 8,60.10^{-5}$$

2) Pour $h = 1/15$ fixé et $p = 4p_i$ nous obtenons à partir de $n = 20$ itérations, l'erreur suivante :

$$\|u_h^n - u\|_{L^2}^2 = 6,34.10^{-5}$$

Pour les deux cas le temps d'exécution en C.D.C. 3 600 est de l'ordre de quelques secondes.

Pour être sûr que nous avons effectivement résolu l'inéquation (3.13) sur le convexe X donné par (3.36); en effectuant les itérations (3.25) nous avons fait

$$f_h = f(x, y) \text{ quelconque } (= e^{xy} \text{ pour l'expérience})$$

et on vérifie que la solution u_h^n obtenue après stationarité des résultats appartient bien au convexe X .

III-4 Problème 4

Un exemple unidimensionnel : X est un convexe sur lequel on ne sait pas projeter directement.

On cherche $u \in X$ convexe fermé de $W_0^{1,p}(0,1)$ vérifiant :

$$(3.38) \quad (A, v - u) \geq (f, v - u) \quad \forall v \in X \quad f \in W^{-1,q}(0,1)$$

où

$$(3.39) \quad A u = \frac{d}{dx} \left(\left| \frac{du}{dx} \right|^{p-2} \frac{du}{dx} \right) - \lambda \frac{d^2 u}{dx^2} \quad \lambda > 0 ; p \geq 2$$

$$(3.40) \quad X = \{ u \mid u \in W_0^{1,p}(0,1) ; \left| \frac{du}{dx} \right| \leq 1 \text{ p.p sur } (0,1) \}$$

Soient m un entier et h un paramètre positif tel que $mh = 1$.

On divise l'intervalle $(0,1)$ en m intervalles : $(0,h)$, $(h, 2h)$, ..., $((m-1)h, mh)$. Soit θ_h^i la fonction caractéristique de $]ih, (i+1)h[$ pour $i = 0, \dots, m-1$.

$$V_h = \{ u_h \mid u_h = \sum_{i=0}^{m-1} u_h^i \theta_h^i, u_h^i \in \mathbb{R} ; u_h^0 = u_h^m = 0 \}$$

Posons

$$\nabla u_h(x) = \frac{u_h(x + \frac{h}{2}) - u_h(x - \frac{h}{2})}{h}$$

$$\nabla^2 u_h = \nabla (\nabla u_h)$$

$$\text{Enfin} \quad \hat{\nabla} u_h(x) = \frac{u_h(x) - u_h(x-h)}{h}$$

Le problème (3.38), (3.39), (3.40) peut alors être approché par

$$(3.41) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h \quad \text{où}$$

$$(3.42) \quad A_h u_h = -\nabla \left(\left| \nabla u_h \right|^{p-2} \nabla u_h \right) - \lambda \nabla^2 u_h \quad p \geq 2, \quad \lambda > 0.$$

$$(3.43) \quad X_h = \{ u_h \mid u_h \in V_h ; \left| \hat{\nabla} u_h \right| \leq 1 \text{ p.p} \}$$

Nous allons résoudre numériquement (3.41), (3.42), (3.43) par deux méthodes.

1ère méthode

Posons

$S_h u_h = -\nabla^2 u_h$. On munit V_h du nouveau produit scalaire

$$[u_h, v_h] = (S_h u_h, v_h)_h = (\nabla u_h, \nabla v_h)_{L^2}$$

et de la norme associée $[u_h] = \|\nabla u_h\|_{L^2}^2$.

On a alors la

Proposition 3.2

1 - La suite (u_h^n) définie par les itérations :

$$(3.44) \quad u_h^{n+1} = P_{X_h}^{S_h} (u_h^n - \rho S_h^{-1} (A_h u_h^n - f_h))$$

où $S_h = -\nabla^2$ converge dans V_h vers u_h solution du problème

(3.41), (3.42), (3.43) pour

$$u_h^0 = 0, \quad \rho = \lambda / C^2(N) \quad \text{où} \quad N = \frac{2}{\lambda} \|\mathcal{J}_h\|_{L^2}^2$$

$$\text{et} \quad C(N) = (p-1) \frac{N^{p-2}}{h^{\frac{p-2}{2}}} + \lambda; \quad p \geq 2$$

2 - De plus on a $P_h u_h \longrightarrow u$ dans $W_0^{1,p}(\Omega)$ fort quand $h \longrightarrow 0$, où u est la solution du problème (3.38), (3.39), (3.40).

Pour que les itérations (3.44) soient possibles numériquement ; il est nécessaire de savoir projeter explicitement sur le convexe X_h donné par (3.43).. On se propose maintenant de résoudre ce point.

Notations

Comme on est dans V_h espace de dimension finie on écrira dorénavant

$$u_h(x) = u(i) \quad i = 1, \dots, m$$

$$f_h(x) = f(i) \quad i = 1, \dots, m$$

Soit T_h une application de V_h dans V_h' définie par

$$T_h u(t) = \sum_{j=1}^i h u(j) \quad (1)$$

Nous avons alors

$$T_h \hat{V} u(t) = \hat{V} T_h u(t) = u(t) \quad \forall t$$

Donc T_h est un isomorphisme de V_h dans V_h' d'inverse

$$T_h^{-1} = \hat{V}$$

Il s'agit maintenant de calculer explicitement la projection sur

$$X_h \text{ de } f_n(t) = u_n(t) - \rho S_h^{-1} (A_h u_n(t) - f(t)).$$

On va projeter sur le convexe $\tilde{X}_h$ et on se ramènera ensuite au convexe X_h . Plus précisément, posons :

$$\tilde{X}_h = T_h^{-1}(X_h) = \{ \tilde{w} \mid |\tilde{w}(t)| \leq 1 \text{ pour } t = 1, \dots, m \}$$

$$\tilde{B}_h = T_h^* S_h T_h = (\hat{V}_\epsilon^{-1})^* (-V_\epsilon^L) \hat{V}_\epsilon^{-1} \quad (1)$$

$$\tilde{f}_n = T_h^* S_h f_n$$

Pour n fixé le calcul de

$$v_n = P_{X_h}^{S_h}(f_n) \text{ équivaut à résoudre}$$

$$(S_h v_n, w - v_n) \geq (S_h f_n, w - v_n) \quad \forall w \in X_h$$

Posons $v_n = T \tilde{v}_n$ $w = T \tilde{w}$. Il vient alors

$$(3.45) \quad (\tilde{B}_h \tilde{v}_n, \tilde{w} - \tilde{v}_n) \geq (\tilde{f}_n, \tilde{w} - \tilde{v}_n) \quad \forall \tilde{w} \in \tilde{X}_h$$

(1) C'est une intégration discrète. Il y a d'autres formules bien plus précises.

Plus précisément l'inverse de $\hat{V}$ sera $\hat{V}_\epsilon^{-1} = (\hat{V} - \epsilon \hat{V}^2)^{-1}$ avec $\epsilon \rightarrow 0, \epsilon > 0$.

Comme $\tilde{B} \tilde{v}_n \supset \tilde{v}_n$ pour h suffisamment petit la relation (3.45) s'écrit alors

$$(3.46) \quad \tilde{v}_n = P_{\tilde{X}_h}(\hat{v} f_n)$$

on sait donc par (3.46) résoudre la projection sur $\tilde{X}_h$

$$\text{où } \tilde{X}_h = \{ \tilde{w} \mid |\tilde{w}(t)| \leq 1 \text{ p.p.} \}$$

On peut maintenant combiner les deux formules itératives (3.44) et (3.46) en une seule.

Proposition 3.3

1 - La suite (u_h^n) définie par les itérations

$$(3.47) \quad u_h^{n+1} = T [P_{\tilde{X}}(\hat{v} f_h^n)] \quad \text{où } T = \hat{v}^{-1} = (\hat{v} - \epsilon \nabla^2)^{-1} \epsilon > 0 \text{ et } \epsilon \rightarrow 0.$$

$$\text{avec } f_h^n = u_h^n - \rho S_h^{-1} (A_h u_h^n - f_h) \quad ; \quad S_h = -\nabla^2$$

Converge dans V_h vers u_h solution du problème (3.41), (3.42), (3.43) pour

$$u_h^0 = 0, \quad \rho = \lambda / C^2(N) \quad \text{où } C(N) = (p-1) \frac{N^{p-2}}{h^{\frac{p-2}{2}}} + \lambda \quad p \geq 2$$

$$\text{et } N = \frac{2}{\lambda} \|f_h\|_L^2$$

2 - De plus, on a $p_h u_h \longrightarrow u$ dans $W_0^{1,p}(0,1)$ fort quand $h \longrightarrow 0$; u étant la solution du problème (3.38), (3.39), (3.40).

Remarque 3.10

Pour n fixé (3.46) s'écrit, pour $i = 1, \dots, m$

$$v_n(i) = \begin{cases} \hat{v} f_n(i) & \text{si } -1 \leq \hat{v} f_n(i) \leq 1 \\ -1 & \text{si } \hat{v} f_n(i) < -1 \\ 1 & \text{si } \hat{v} f_n(i) > 1 \end{cases}$$

2ème méthode

Posons comme dans la 1ère méthode

$$T_h u(i) = \sum_{j=1}^l h u(j) \quad i = 1, \dots, m \quad \text{ou encore} \quad T_h u = \nabla_\epsilon^{-1} u$$

Posons $\tilde{X}_h = \{ \tilde{v} \mid |\tilde{v}(i)| \leq 1 \text{ pour } i = 1, \dots, m \}$

Le problème (3.41), (3.42), (3.43) équivaut alors à résoudre

$$(3.48) \quad (\tilde{A}_h \tilde{u}_h, \tilde{v}_h - \tilde{u}_h) \geq (\tilde{f}_h, \tilde{v}_h - \tilde{u}_h) \quad \forall \tilde{v}_h \in \tilde{X}_h \quad \text{où}$$

$$\begin{aligned} \tilde{A}_h \tilde{u}_h &= T_h^* A_h T_h \tilde{u}_h = \hat{\nabla}_\epsilon^{-1} \nabla (|\nabla \hat{\nabla}_\epsilon^{-1} \tilde{u}_h|^{p-2} \nabla \hat{\nabla}_\epsilon^{-1} \tilde{u}_h) \\ &\quad + \lambda \hat{\nabla}_\epsilon^{-1} \nabla^2 \hat{\nabla}_\epsilon^{-1} \tilde{u}_h \quad \lambda > 0 \end{aligned}$$

$$\tilde{f}_h = T_h^* f_h$$

et (3.48) se résoud alors par itérations.

Proposition 3.4

1 - La suite $(\tilde{u}_h^n)$ définie par les itérations

$$(3.49) \quad \tilde{u}_h^{n+1} = P_{\tilde{X}_h} (\tilde{u}_h^n - \tilde{\rho} (\tilde{A}_h \tilde{u}_h^n - \tilde{f}_h))$$

et convergente pour $\tilde{u}_h^0 = 0$ et $\tilde{\rho} > 0$ convenable

2 - La solution u_h du problème (3.41), (3.42), (3.43) est donnée par

$$(3.50) \quad u_h = T \left(\lim_{n \rightarrow \infty} \tilde{u}_h^n \right)$$

3 - De plus on a $u = \lim_{h \rightarrow 0} p_h u_h$ dans $W_0^{1,p}(0,1)$ fort où u est la solution du problème (3.38), (3.39), (3.40).

Remarque 3.11

Si on travaille avec un opérateur de prolongement p_h de V_h dans V conduisant à l'inéquation

$$(3.51) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h) \quad \forall v_h \in X_h$$

avec $A_h u_h = - \hat{\nabla} (|\hat{\nabla} u_h|^{p-2} \hat{\nabla} u_h) - \lambda \hat{\nabla}^2 u_h$

et $X_h = \{ u_h \in V_h \mid |\hat{\nabla} u_h| \leq 1 \text{ p.p.} \}$

Alors le choix $T_h = \hat{\nabla}_\epsilon^{-1}$ conduit à l'inégalité

$$(3.52) \quad (\tilde{A}_h \tilde{u}_h, \tilde{v}_h - \tilde{u}_h)_h \geq (\tilde{f}_h, \tilde{v}_h - \tilde{u}_h)_h \quad \forall \tilde{v}_h \in \tilde{X}_h$$

avec $\tilde{A}_h \tilde{u}_h = T_h^* A_h T_h \tilde{u}_h \simeq |\tilde{u}_h|^{p-2} \tilde{u}_h + \lambda \tilde{u}_h \quad \lambda > 0$

$$\tilde{f}_h(i) = T_h^* f_h(i) \quad i = 1, \dots, m$$

$$\tilde{X}_h = \{ \tilde{u}_h \mid |\tilde{u}_h| \leq 1 \text{ p.p.} \}$$

L'inégalité (3.52) peut alors se résoudre par la formule itérative suivante :

$$(3.53) \quad \tilde{u}_h^{n+1} = P_{\tilde{X}} (\tilde{u}_h^n - \tilde{\rho} (\tilde{A}_h \tilde{u}_h^n - \tilde{f}_h))$$

qui converge pour $\tilde{u}_h^0 = 0$ et $\tilde{\rho} > 0$ qu'il est maintenant aisé de calculer. En effet :

$$(\tilde{A}_h \tilde{u}_h^n - \tilde{A}_h \tilde{v}_h, \tilde{u}_h^n - \tilde{v}_h)_h \geq \lambda \|\tilde{u}_h^n - \tilde{v}_h\|_L^2$$

Si $\|\tilde{u}_h\| \leq N$ et $\|\tilde{v}_h\| \leq N$ on a :

$$(\tilde{A}_h \tilde{u}_h^n - \tilde{A}_h \tilde{v}_h, \tilde{u}_h^n - \tilde{v}_h) \leq C(N) \|\tilde{u}_h^n - \tilde{v}_h\| \|\tilde{u}_h^n\| \quad \text{avec } C(N) = (p-1) N^{p-2} + \lambda$$

d'où $\rho = \lambda / C^2(N)$.

Résultats numériques.

On veut résoudre l'inéquation :

$$(3.54) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X \text{ pour } f \text{ donné dans } L^{3/2}(0,1)$$

avec

$$Au = - \frac{d}{dx} \left(\left| \frac{du}{dx} \right| \frac{du}{dx} \right) - \lambda \frac{d^2 u}{dx^2} \quad \lambda = 10$$

$$X = \left\{ u \in W_0^{1,3}(0,1) \mid \left| \frac{du}{dx} \right| \leq 1 \text{ p.p.} \right\}$$

Au problème (3.54) on associe le problème

$$(3.55) \quad (A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

avec

$$A_h u_h = - \nabla (|\nabla u_h| \nabla u_h) - \lambda \nabla^2 u_h$$

$$X_h = \left\{ u_h \in V_h \mid |\hat{\nabla} u_h| \leq 1 \text{ p.p.} \right\}$$

Ce problème est alors résolu à l'aide des itérations

$$(3.56) \quad u_h^{n+1} = \hat{V}_\epsilon^{-1} P_{\hat{X}_h} (\hat{V} f_h^n) \quad \text{avec } u_h^0 = 100, \quad \epsilon \longrightarrow 0$$

$$f_h^n = u_h^n - \rho S_h^{-1} (A_h u_h^n - f_h) ; \quad S_h = - \nabla^2$$

$$\rho = \lambda / C^2(N) \quad \text{et} \quad C(N) = \frac{4}{\lambda \sqrt{h}} \|f_h\|_{L^2} + \lambda$$

$$\hat{X}_h = \left\{ \tilde{w}_h \mid |\tilde{w}_h| \leq 1 \text{ p.p.} \right\}$$

Plus précisément nous avons procédé de la manière suivante :

1) On se donne une solution exacte

$$(3.57) \quad u = x - x^2 \quad x \in (0,1)$$

et on pose

$$f = Au$$

ce qui donne

$$P_X^S (u - \rho S^{-1} (Au - f)) = u \quad ; \quad \text{car } u \in X$$

La formule itérative (3.56) donne alors une solution (u_h^n) que l'on compare u donné par (3.57).

2) Pour bien voir si le programme de projection sur X_h fonctionne bien, nous avons donné comme première valeur initiale $u_h^0 = 100$ et il y a alors nécessité de projeter.

3) Il est essentiel pour l'exemple (3.54) de connaître la région de l'intervalle $(0,1)$ sur laquelle on a $|\frac{du}{dx}| = 1$. Pour cela on se donne un second membre f ($f = \exp(8x^3)$ pour l'expérience) ; on calcule u_h^n par itérations jusqu'à stabilisation des résultats.

Pour n assez grand (cf tableau ci-dessous) on a :

$$\begin{aligned} |\text{grad } u_h^n| &< 1 \quad \text{pour } 0 \leq x \leq 0,7 \\ \text{et } |\text{grad } u_h^n| &= 1 \quad \text{pour } 0,7 \leq x \leq 1 \end{aligned}$$

Nous obtenons les résultats numériques suivants : (sur C.D.C 3.600)

	$h = 1/10$	$h = 1/20$	$h = 1/40$	$h = 1/40$
Nombre d'itérations	300	700	1 000	2 000
Temps d'exécution	4,7 s	30 s	3 mn 47 s	4 mn 21 s
Erreur $\ u_h^n - u\ _{L^2}$	$1,95 \cdot 10^{-3}$	$9,85 \cdot 10^{-4}$	$7,86 \cdot 10^{-4}$	$6,31 \cdot 10^{-4}$

Remarque 3.12

Dans le tableau ci dessus on remarque que pour $h = 1/40$ pour faire 1 000 itérations il a fallu 3 mn 47 s et pour effectuer 2 000 itérations seulement 4 mn 21 s . Cela tient à ce qu'au bout d'un certain nombre d'itérations, la solution u_h^n est dans le convexe et le programme de projection a été évité.

Les courbes n° 4 donnent l'évolution de l'erreur $\|u_h^n - u\|_{L^2}$ en fonction du nombre n d'itérations et pour $h = 1/10$, $h = 1/20$, $h = 1/40$.

Remarque 3.13'

Quand on effectue les itérations en prenant $\underline{\rho = 4\rho_0}$, ρ_0 étant la valeur théorique, on a une accélération très nette de la convergence :

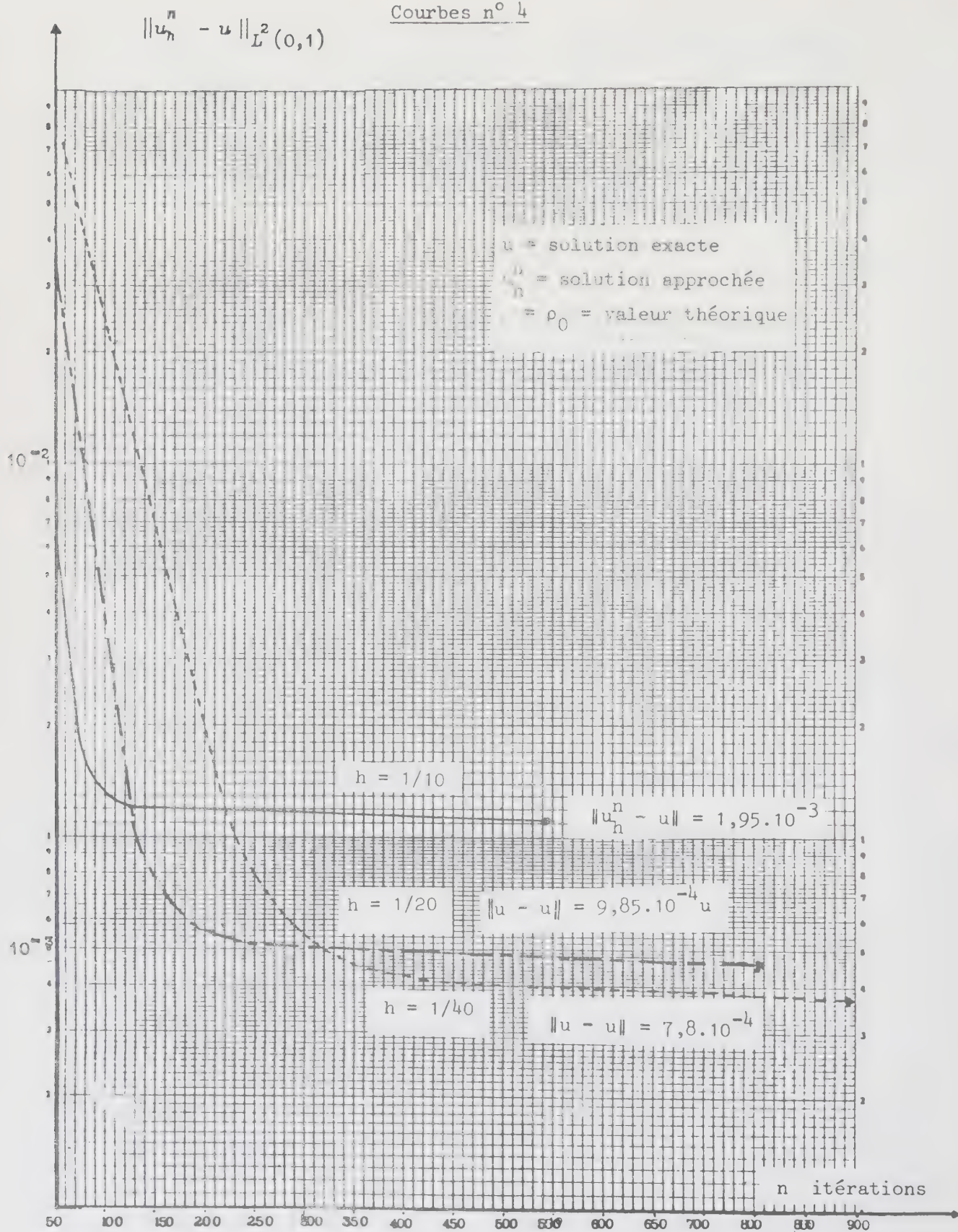
Pour $h = 1/20$ et $n = 50$ on a $\|u_h^n - u\|_{L^2} = 9.10^{-4}$

Pour $h = 1/40$ et $n = 70$ on a $\|u_h^n - u\|_{L^2} = 2.10^{-4}$

(résultats à comparer avec le tableau ci-dessus).

Le temps d'exécution est inférieur à la minute (exécution + compilation).

Courbes n° 4



III-5 Un problème n-dimensionnel : X est un convexe sur lequel on ne sait pas projeter explicitement.

Soit $\mathcal{H}$ un espace de Hilbert sur $\mathbb{R}^n$ et Ω un ouvert borné de $\mathbb{R}^n$ de frontière $\Gamma = \partial \Omega$ suffisamment régulière. On se donne un convexe fermé X de $\mathcal{H}$ défini par

$$(3.58) \quad X = \{ u \in \mathcal{H} \mid |\text{grad } u| \leq 1 \text{ p.p. sur } \Omega \}$$

Problème 1

On cherche $u \in X$ tel que

$$(3.59) \quad (Au, v - u) \geq (f, v - u) \quad \forall v \in X$$

pour f donné dans $\mathcal{H}'$ et A est un opérateur non nécessairement linéaire vérifiant les hypothèses H2) du théorème 1.1 § I.

Problème 2

On se donne un opérateur $S \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ vérifiant les hypothèses H1) du théorème 1.1 § I.

On cherche $u \in X$ tel que

$$(3.60) \quad (Su, v - u) \geq (g, v - u) \quad \forall v \in X$$

pour g donné dans $\mathcal{H}'$ et X donné par (3.58).

Problème 2 et projection sur X suivant une norme associée à S

Le théorème 1.1 §I dit que la suite (u_n) définie par

$$(3.61) \quad u_{n+1} = P_X^S (u_n - \rho S^{-1} (A u_n - f))$$

converge fortement dans $\mathcal{H}$ vers u solution de (3.59). Mais la relation (3.61) suppose que l'on sache calculer explicitement

$P_X^S(w)$ pour $w \notin X$; La projection sur X donné par (3.58)

étant prise au sens du produit scalaire $[u, v] = (S u, v)$.

Nous allons maintenant chercher à résoudre (3.60) pour le convexe X donné par (3.58) et pour un certain opérateur S . La résolution de (3.59) ne posera ensuite aucun problème.

Proposition 3.5

- 1 - Le problème (3.60) admet une solution unique $u \in X$
- 2 - Toute solution de (3.60) est solution de

$$(3.62) \quad u = P_X^S (S^{-1} g)$$

la projection étant prise au sens de la norme $[u]^2 = (S u, u)$ et réciproquement.

Démonstration

- 1 - Soit $S \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ et vérifiant

$$(S u, v) = (u, S v) \quad \forall u, v \in \mathcal{H}.$$

On munit alors $\mathcal{K}$ du nouveau produit scalaire

$$[u, v] = (S u, v) \quad \text{et de la norme} \quad [u]^2 = (S u, u)$$

On suppose qu'il existe une constante $\sigma > 0$ telle que

$$[u]^2 \geq \sigma \|u\|^2 \quad \text{où} \quad \| \cdot \| \quad \text{est la norme de} \quad \mathcal{K} \quad \text{associée au produit scalaire} \quad (,) .$$

Il suffit alors d'appliquer le théorème 1.1 § I.

2 - (3.60) est équivalent à (3.62) par définition de la projection sur X suivant le produit scalaire $[u, v]$.

Discrétisation

Soit $h = (h_1, \dots, h_n)$ un paramètre positif de R^n destiné à tendre vers 0. On suppose qu'il existe une famille d'approximations de l'espace $\mathcal{K}$; $\{\mathcal{K}, V_h, p_h, r_h\}$ stable et consistante, cf [1] où V_h est un espace de dimension finie. p_h est un prolongement linéaire continu de V_h dans $\mathcal{K}$ et r_h est une restriction linéaire continue de $\mathcal{K}$ dans V_h . On note par $\hat{r}_h$ la transposée de p_h . Nous avons alors le schéma suivant

$$\begin{array}{ccc} \mathcal{K} & \xrightarrow{S} & \mathcal{K}' \\ p \uparrow & & \downarrow \hat{r}_h \\ & r_h & \\ V_h & \xrightarrow{\quad} & V_h' \end{array}$$

On munit V_h des produits scalaires $(,)_h$ et $[,]_h$ et des normes associées $\| \cdot \|_h$ et $[\cdot]_h$.

Posons

$$\hat{\nabla}_i u_h(x) = \frac{u_h(x) - u_h(x - h_i)}{h_i} \quad i = 1, \dots, n \quad \forall u_h \in V_h$$

Le problème (3.58), (3.60) s'écrit alors en dimension finie sous la forme :

$$(3.63) \quad (S_h u_h, v_h - u_h)_h \geq (\theta_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

$$(3.64) \quad X_h = \{ u_h \in V_h \mid \sum_{i=1}^n |\hat{\nabla}_i u_h|^2 \leq 1 \text{ p.p.} \}$$

Remarque 3.14

Si u_h est solution de (3.63), (3.64) la question de la convergence de $p_h u_h$ vers u solution de (3.58), (3.60) ne pose aucun problème. Il est néanmoins utile de discrétiser le problème car sur un ordinateur, on ne peut travailler qu'en dimension finie.

Cas où Ω est un ouvert de R^2

Nous allons résoudre numériquement en dimension $n = 2$ les problèmes (3.63), (3.64). Par commodité, on supprime l'indice h . On posera en particulier :

$$u_h(x) = u(i, j) \quad i = 1, \dots, l_1 \quad ; \quad j = 1, \dots, l_2$$

Le couple (i, j) désignant alors le point générique d'un réseau R_h de R^2 associé au pas h .

Proposition 3.6

Soit $\theta : V \longrightarrow V \times V$ défini par

$$(3.65) \quad \theta u = \frac{1}{2} (\hat{\nabla}_1 u, \hat{\nabla}_2 u)$$

Soit $T : V \times V \longrightarrow V$ défini par

$$(3.66) \quad T(u_1, u_2) = A_1^{-1} u_1 + A_2^{-1} u_2$$

$$\text{où } A_1 = -\epsilon \hat{\nabla}_1^2 + \hat{\nabla}_1$$

$$A_2 = -\epsilon \hat{\nabla}_2^2 + \hat{\nabla}_2$$

avec des conditions aux limites Duichlét et ϵ étant un paramètre positif arbitrairement petit destiné à tendre vers 0.

Posons

$$\begin{aligned} W &= \{ \tilde{u} \mid \tilde{u} = (u_1, u_2) \mid \hat{v}_1 u_2 = \hat{v}_2 u_1 \} \\ \tilde{X} &= \{ \tilde{v} \in W \mid |\tilde{u}|^2 = |u_1|^2 + |u_2|^2 \leq 1 \text{ p.p} \} \\ \tilde{S} &= T^* S T \quad \text{où} \quad T^* \text{ est l'adjoint de } T. \end{aligned}$$

On suppose que $\tilde{g} = T^* g \in W$

On munit alors W de la norme

$$[\tilde{u}]_1 = (\tilde{S} \tilde{u}, \tilde{u})^{1/2}$$

Alors la solution du problème (3.63), (3.64) s'écrit:

$$(3.67) \quad u = T \tilde{u} = T P_{\tilde{X}} (\theta S^{-1} g)$$

où la projection sur $\tilde{X}$ est prise au sens de la norme $[\tilde{u}]_1$.

Démonstration

Posons $u = T \tilde{u}$, $v = T \tilde{v}$ dans (3.63); il vient alors

$$(S T \tilde{u}, T \tilde{v} - T \tilde{u}) \geq (T^* g, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X}$$

Donc

$$(3.68) \quad (\tilde{S} \tilde{u}, \tilde{v} - \tilde{u}) \geq (\tilde{g}, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X}$$

où $\tilde{S} = T^* S T$, $\tilde{g} = T^* g$

Si on munit W de la norme $[\tilde{u}]_1 = (\tilde{S} \tilde{u}, \tilde{u})^{1/2}$ d'après la proposition 3.5 le problème (3.68) est équivalent à :

(3.69) $\tilde{u} = P_{\tilde{X}} (\tilde{S}^{-1} \tilde{g})$ où la projection sur $\tilde{X}$ est prise au sens de la norme $[\tilde{u}]_1$.

Par ailleurs, on a :

$$T \theta u = T \frac{1}{2} (\hat{V}_1 u, \hat{V}_2 u) = u$$

$$\text{et} \quad \tilde{S}^{-1} \tilde{g} = (T^* S T)^{-1} T^* g = \theta S^{-1} g$$

$$\text{car} \quad (T^* S T) \theta g = T^* g \quad . \quad \text{D'où} \quad (3.67).$$

En général ce que l'on sait faire explicitement, c'est projeter sur le convexe

$$\tilde{X} = \{ \tilde{u} \in W \mid |\tilde{u}| \leq 1 \text{ p.p} \}$$

suivant la norme usuelle $|\tilde{u}|$.

Proposition 3.7

On se place sous les hypothèses de la proposition

Si de plus on a $\tilde{S} = T^* S T = I$ d, alors on peut calculer explicitement la solution u du problème

Plus précisément posons :

$$\varphi_{t,j} = |\tilde{u}|^2 = |u_1(t, j)|^2 + |u_2(t, j)|^2$$

$$\tilde{g} = T^* g = (g_1, g_2)$$

$$(3.70) \quad \tilde{u}(t, j) \begin{cases} (g_1(t, j), g_2(t, j)) & \text{si } \varphi_{t,j} \leq 1 \\ \left(\frac{g_1(t, j)}{\sqrt{\varphi_{t,j}}}, \frac{g_2(t, j)}{\sqrt{\varphi_{t,j}}} \right) & \text{si } \varphi_{t,j} > 1 \end{cases}$$

Alors

$$(3.71) \quad u(t, j) = T \tilde{u}(t, j) = A_1^{-1} u_1(t, j) + A_2^{-1} u_2(t, j)$$

est la solution du problème (3.63, (3.64)).

Démonstration

Si $\tilde{S} = T^* S T = I d$ alors (3.68) devient :

$$(\tilde{u}, \tilde{v} - \tilde{u}) \geq (\tilde{g}, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X}$$

ce qui équivaut à

$$(3.72) \quad \tilde{u} = P_{\tilde{X}}(\tilde{g})$$

où la projection est prise cette fois au sens de la norme usuelle

$$|\tilde{u}|^2 = |u_1|^2 + |u_2|^2.$$

Pour (i, j) fixé, le convexe $\tilde{X}$ est un disque et l'on sait projeter explicitement sur $\tilde{X}$ suivant la norme $|\tilde{u}|$. D'où les formules (3.70), (3.71).

Si $\tilde{S} = T^* S T \neq I d$, on peut seulement calculer la solution de (3.63), (3.64) avec des itérations.

Plus précisément, on a la

Proposition 3.8

On pose $\tilde{S} = T^* S T \neq I d$; $\tilde{g} = T^* g \in W$; alors

1 - La suite $(\tilde{u}_n)$ définie par l'itération

$$(3.73) \quad \tilde{u}_{n+1} = P_{\tilde{X}}(\tilde{u}_n - \tilde{\rho}(\tilde{S}\tilde{u}_n - \tilde{g}))$$

où la projection est prise au sens de la norme $|\tilde{u}|$ est convergente dans W fort pour $\tilde{u}_0 (= 0)$ et $\tilde{\rho} > 0$ convenables.

2 - La solution $u \in X$ de l'inéquation

$$(Su, v - u) \geq (g, v - u) \quad \forall v \in X \quad \text{s'écrit alors}$$

$$(3.74) \quad u = T \left(\lim_{n \rightarrow \infty} u_n \right)$$

Démonstration

On se propose de résoudre (3.68) par itérations :

Nous avons :

$$(\tilde{S} \tilde{u}, \tilde{v}) \leq C |\tilde{u}| |\tilde{v}| \quad \text{et} \quad (\tilde{S} \tilde{u}, \tilde{u}) \geq \kappa |\tilde{u}|^2$$

On pose $\tilde{\rho} = \kappa/C^2$. D'après le théorème 1.1 chap. I la suite $(\tilde{u}_n)$ définie par (3.73) converge fortement dans W vers $\tilde{u}$ solution de (3.68) et par définition on a $u = T \tilde{u}$. D'où (3.74).

Remarque 3.15

Dans les itérations (3.73) le calcul explicite de $\tilde{\rho}$ est difficile, mais on peut toujours prendre $\tilde{\rho} = \tilde{\rho}_n$ et appliquer les remarques du paragraphe I-1.2.

Remarque 3.16

1 - Le fait de supposer dans la proposition 3.7 $\tilde{S} = T^* S T = I$ est raisonnable. Voici un exemple en dimension $n = 1$.

On pose $T = \hat{V}_\epsilon^{-1}$ et $S = -\hat{V}_\epsilon^2 \simeq -\frac{d^2}{dx^2}$; où $\hat{V}_\epsilon^{-1} = (\hat{V} - \epsilon \hat{V}^2)^{-1}$.

2 - Si $S = (T^* T)^{-1}$ d'après la proposition 3.7 on peut alors résoudre explicitement et sans itérations le problème (3.63), (3.64).

Résolution du problème 1 avec un opérateur $A \neq S$

La résolution numérique de l'inéquation :

$$(3.75) \quad (S u, v - u) \geq (g, v - u) \quad \forall v \in X$$

pour $X = \{ u \mid |\text{grad } u| \leq 1 \text{ p.p.} \}$

n'est qu'un moyen pour résoudre l'inéquation plus générale

$$(3.76) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

considérée dans le problème 1.

Nous allons montrer que dans ce cas l'opérateur S de (3.75) est à notre disposition et on pourrait le choisir de manière que $S = T^* S T = I$; ce qui permettra de résoudre (3.76) avec une formule itérative unique.

Proposition 3.9

Soit A un opérateur non nécessairement linéaire de V_h dans V_h' vérifiant les hypothèses du théorème 1.1 chap. I. Alors

1 - Il existe $u \in X = \{u \mid |\text{grad } u| \leq 1 \text{ p.p}\}$ unique solution de

$$(3.76) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

2 - La suite (u_n) définie par les itérations :

$$(3.77) \quad u_{n+1} = P_X^S (u_n - \rho S^{-1} (A u_n - f))$$

converge dans V_h vers u solution de (3.76) quand $n \rightarrow \infty$

pour $u_0 (= 0)$ et ρ convenables.

3 - Posons $f_n = u_n - \rho S^{-1} (A u_n - f)$. Pour calculer explicitement sur le convexe X

$$u_{n+1} = P_X^S (f_n) \quad \text{il suffit alors de résoudre}$$

$$(3.78) \quad (S u_{n+1}, w - u_{n+1}) \geq (S f_n, w - u_{n+1}) \quad \forall w \in X.$$

Démonstration

Les assertions 1 - et 2 - sont évidentes. Démontrons 3 -.

On doit calculer explicitement

$$u_{n+1} = P_X^S (f_n) \quad \text{la projection étant prise au sens du produit scalaire}$$

$[u, v] = (S u, v)$. Il revient alors au même de résoudre (3.78).

On est alors dans le cadre du problème 2 que l'on vient de résoudre en posant $g = S f_n$ pour n fixé.

Remarque 3.17

Pour passer du convexe X au convexe $\tilde{X}$, les opérateurs T et θ s'imposent. Par contre le choix de l'opérateur S est libre. Si on pose $S = (T T^*)^{-1}$ dans (3.78) on aura :

$$u_{n+1} = T \tilde{u}_{n+1} ; \quad w = T \tilde{w} \quad \text{donc}$$

$$(\tilde{u}_{n+1}, \tilde{w} - \tilde{u}_{n+1}) \geq (\tilde{\theta}_n, \tilde{w} - \tilde{u}_{n+1})$$

avec
$$\tilde{\theta}_n = T^* (T T^*)^{-1} f_n = \theta f_n$$

On a alors

$$\tilde{u}_{n+1} = P_{\tilde{X}}^{\sim}(\theta f_n) \quad \text{D'où la}$$

Proposition 3.10

La suite (u_n) définie par les itérations

$$u_{n+1} = T P_{\tilde{X}}^{\sim}(\theta u_n - \rho T^*(A u_n - f))$$

où
$$\tilde{X} = \{ \tilde{u} \in w \mid |\tilde{u}|^2 = |u_1|^2 + |u_2|^2 \leq 1 \quad p.p \}$$

et où la projection sur $\tilde{X}$ est prise au sens de la norme $|\tilde{u}|$ converge dans V_h fort pour $u_0 = 0$ et $\rho = h/C^2(N)$ vers u_h solution de

$$(A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

où
$$X_h = \{ u_h \in V_h \mid |\nabla_1 u_h|^2 + |\nabla_2 u_h|^2 \leq 1 \quad p.p \}$$

Par ailleurs on a $p_h u_h \longrightarrow u$ dans V fort ; u étant la solution de

$$(A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

où
$$X = \{ u \in V \mid |\text{grad } u| \leq 1 \quad p.p \}$$

Remarque 3.18

On sait que le problème

$$(A_h u_h, v_h - u_h)_h \geq (f_h, v_h - u_h)_h \quad \forall v_h \in X_h$$

$$\text{où } A_h u_h = - \sum_{i=1}^n \nabla_i (|\nabla_i u_h|^{p-2} \nabla_i u_h) - \lambda \sum_{i=1}^n \nabla_i^2 u_h$$

$$X_h = \left\{ u_h \in V_h; \left(\sum_{i=1}^n |\nabla_i u_h|^2 \right)^{1/2} \leq 1 \text{ p.p.} \right\}$$

est équivalent au problème

$$F_h(u_h) \leq F_h(v_h) \quad \forall v_h \in X_h$$

$$\text{où } F_h(u_h) = \frac{1}{p} \sum_{i=1}^n \|\nabla_i u_h\|_{L_p}^p + \frac{\lambda}{2} \sum_{i=1}^n \|\nabla_i u_h\|_{L_2}^2 - (f, u_h)$$

La méthode de la mire que nous avons introduit dans [14] permet de résoudre ce problème directement.

L'utilisation d'une méthode directe (méthode de la mire par exemple) semble s'imposer chaque fois que le convexe X_h est très "compliqué" (cf. les résultats numériques dans [14]).

III-6 Résolution d'un problème de mécanique

Soit à résoudre l'inéquation

$$(3.79) \quad (A u, v - u) \geq (f, v - u) \quad \forall v \in X$$

$$\text{où } X = \left\{ v \in H_0^1(\Omega) ; |\text{grad } u| \leq 1 \text{ p.p. sur } \Omega \right\}$$

Ω étant un ouvert borné de R^n .

avec
$$A = -\Delta = - \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$$

et $f = c^{te}$ donné.

L'inégalité (3.79) est le problème physique correspondant à la torsion élastoplastique (cf. [4], [14], [15]). C'est en outre un cas particulier du problème (3.59).

On se propose maintenant de résoudre numériquement (3.79) par une méthode simple.

Plus précisément, nous avons le

Théorème 3.3 (cf H. Brezis et M. Sibony [4])

1 - Le problème (3.79) équivaut à trouver $\tilde{u} \in \tilde{X}$ solution de

$$(3.80) \quad (A \tilde{u}, \tilde{v} - \tilde{u}) \geq (f, \tilde{v} - \tilde{u}) \quad \forall \tilde{v} \in \tilde{X}$$

$$\text{où } \tilde{X} = \{ \tilde{v} \in H_0^1(\Omega) \mid -\delta \leq v \leq \delta \text{ sur } \Omega \}$$

$$A = -\Delta$$

$$\delta(x) = \text{distance}(x, \partial\Omega)$$

2 - Les problèmes (3.79) et (3.80) admettent une solution unique et on a $u = \tilde{u}$

3 - La solution u de (3.79) est alors donnée par

$$(3.81) \quad u = P_X^A (A^{-1} f)$$

Remarque 3.19

Le calcul numérique de $\tilde{u}$ solution de (3.80) peut se faire de plusieurs manières.

1 - On calcule $v = A^{-1}f$ par la méthode de Gauss-Seidel [16] et on projette v sur $\tilde{X}$ suivant la norme [] ce qui donne u suivant la formule (3.81). Mais en général on ne sait pas projeter numériquement suivant la norme

$$[u] = \left(\sum_{i=1}^n \|D_i u\|_{L_2}^2 \right)^{1/2}.$$

2 - On applique la méthode de la mire à la fonctionnelle

$$F(v) = \frac{1}{2} \sum_{i=1}^n \|D_i v\|_{L_2}^2 - (f, \tilde{v})$$

avec les contraintes $\tilde{v} \in \tilde{X} = \{ \tilde{v} \in H_0^1(\Omega) ; -\delta \leq \tilde{v} \leq \delta \text{ sur } \Omega \}$
Cf résultats numériques dans [14] paragraphe II-4.

3 - On peut aussi utiliser une méthode itérative

$$u_{n+1} = P_{\tilde{X}}(\tilde{u}_n - \rho_n (A \tilde{u}_n - f))$$

avec $A = -\Delta$ et $\tilde{X} = \{ \tilde{u} \in H_0^1(\Omega) ; |\tilde{u}| \leq \delta \text{ sur } \Omega \}$

La projection sur $\tilde{X}$ étant au sens de la norme $\|\cdot\|_{L_2}$.

On montre alors que $\tilde{u}_n \xrightarrow[n \rightarrow \infty]{} \tilde{u}$ solution de (3.80) dans $H_0^1(\Omega)$ fort. (voir les résultats numériques).

4 - On combine 1) et 3) dans les itérations

$$\tilde{u}_{n+1} = P_{\tilde{X}}^{S_i^{-1}}(u_n - \rho_i S_i^{-1}(A \tilde{u}_n - f))$$

en prenant alternativement

$S_i = S_1 = A$ et $\rho_i = \rho_1 = 1$ si $f_n = u_n - \rho_1 S_1^{-1}(A u_n - f) \in X$

et $S_i = S_2 = Id$ et $\rho_i = \rho_2$ si $f_n \notin X$

(Voir proposition 1.2 chap. I).

Résultats numériques

1) Si f est tel que $A^{-1} f \in X$ la solution u de (3.79) est alors donné par (3.81) qui s'écrit alors sous la forme

$$(3.82) \quad u = A^{-1} f$$

L'opérateur $A = -\Delta$ peut alors être inversé à l'aide de la méthode bien connue de relaxationssuccessivescf. [16] ou bien à l'aide de la méthode itérative suivante :

$$(3.83) \quad u_{n+1} = u_n - \rho_n (A u_n - f)$$

$$\text{avec } u_0 = 0 \quad \text{et} \quad \rho_n = \frac{(A v_n, v_n)}{\|A v_n\|^2} \quad \text{où} \quad v_n = u_n - u_{n-1}$$

2) Si $A^{-1} f \in X$ alors la solution de (3.79) est calculée à l'aide des itérations :

$$(3.84) \quad u_{n+1} = P_X(u_n - \rho_n (A u_n - f))$$

$$\text{avec } u_0 = 0 \quad \text{et} \quad \rho_n = \frac{(A v_n, v_n)}{\|A v_n\|^2} \quad \text{où} \quad v_n = u_n - u_{n-1}.$$

Pour $n \leq 5$ nous prenons un paramètre $\rho_n = \rho = \frac{2+\lambda}{(8/h^2 + \lambda)^2}$ calculé

comme dans les hypothèses du théorème 1.1 chap. I.

Nous procédons alors de la manière suivante :

On se donne un paramètre $h_1 = 1/10$ de discrétisation (ce qui correspond à un réseau de 121 ^{points} associé à l'ouvert $\Omega = (0,1) \times (0,1)$). On calcule $u_{h_1}^n$ à l'aide des itérations (3.84) jusqu'à stabilisation des résultats.

Ensuite on fait $h_2 = h_1/2 = 1/20$ (ce qui correspond à un réseau de 441 points) on pose $u_{h_2}^0 = \tilde{u}_{h_1}^n$ ($\tilde{u}_{h_1}^n$ étant l'interpolé de $u_{h_1}^n$ pour un réseau double) et on calcule alors $u_{h_2}^n$ à l'aide (3.84).

Puis on recommence le procédé avec $h_3 = h_2/2$; ce qui donne la solution finale $u_{h_3}^n$ pour un réseau associé à l'ouvert Ω de 1.681 points.

Le tableau 6a résume les expériences faites.

Tableau n° 6a

	Nombre d'itérations pour $h_1 = 1/10$	Nombre d'itérations pour $h_2 = 1/20$	Nombre d'itérations pour $h_3 = 1/40$
$f = 5$	70	150	530
$f = 10$	30	70	210
$f = 15$	10	50	110

Les tableaux 6b, 6c, 6d donnent les répartitions des contraintes pour $h = 1/10$, $h = 1/20$, $h = 1/40$. On a noté (1) le point (i,j) de Ω où $u(i,j) = \delta = \text{distance}(x, \partial \Omega)$.

Ces points coïncident à l'intérieur du domaine Ω avec ceux pour lesquels $|\text{grad } u| = 1$.

La durée des calculs en C.D.C 3600 est de l'ordre de 1mm pour 100 itérations.

Tableau n° 6b

This image shows a single sheet of white paper with horizontal ruling lines. Each line is composed of a series of small black dots spaced evenly apart. The dots are arranged in a grid-like pattern across the entire page, creating a series of horizontal rows. There are no vertical lines or other markings on the paper.

B I B L I O G R A P H I E

- [1] Aubin J.P.
Approximation des espaces de distributions et des opérateurs différentiels. Bull. Société Math. France mémoire 12, 1967, 139 p.
- [2] Aubin J.P.
Behaviour of the error of the approximate solutions of boundary value problems for linear elliptic operators by Galerkin's and finite difference methods. Annali della Scuola Norm. Pisa. (3) 21 (1967), 599-637
- [3] Brezis H. et Sibony M.
Méthodes d'approximation et d'itération pour les opérateurs monotones. Archive for Rat. Meca and Analysis Vol 28 n° 1 1968 p. 59-82
- [4] Brezis H. et Sibony M.
Sur l'équivalence de certaines inéquations variationnelles (A paraître dans Archive for Rat. Meca and Analysis)
- [5] Brezis H. et F. Stampacchia
Sur la régularité de la solution d'inéquations elliptiques Bull. Soc. Math. France 96, 1968 p. 153-180
- [6] Cea J.
Approximation variationnelle des problèmes aux limites. Annales Inst. Fourier 14 (1964) p. 345-444
- [7] Ciarlet P.G.
Variational methods for non linear boundary value problems (Thèse Case Institute of Technology 1966)
- [8] Goldstein A.A
Minimizing functionals on Hilbert Spaces Computing methods in optimisation problems. Acad. Press 1964
- [9] Lions J.L.
Problèmes aux limites dans les équations aux dérivées partielles Université de Montréal 1962
- [10] Lions J.L. et Stampacchia G.
Variational inequalities. Comm. Pure and Applied Math. Sept. 1967. t. 20, p. 493-519

[10'] Lions J.L. et R. Temam

Une méthode d'éclatement des opérateurs et des contraintes en calcul des variations. C.R. Acad. Sc. Paris, t. 263 (24 Octobre 1966) série A 563-564

[11] Mosco U.

A remark on a theorem of F.E. Browder. Journal of Math. Analysis and Applications Vol 20 n° 1, Oct. 1967

[12] Petryshin W.V.

On the extension and the solution of non linear operator equations. III. Journal of Math. Vol. 10, n° 2 June 1966

[13] Sibony M.

Minimisation de fonctionnelles non différentiables. (A paraître dans Israël Journal of Mathematics)

[14] Sibony M.

Sur l'approximation d'équations et inéquations aux dérivées partielles non linéaires de type monotone (A paraître dans Journal of Mathematical Analysis and Applications)

[14'] Temam R.

Sur la stabilité et la convergence de la méthode des pas fractionnaires. Thèse Fac. Sciences de Paris 1967

[15] Tink

Elastic Plastic Rorsion. Arch. Mech. Anal. t. 25 (1967)

[16] Varga R.S.

Matrix iterative analysis. Englewood Cliffs (Prentice Hall Series in automatic computation). 1962

[17] Yanenko N.N.

Méthode des pas fractionnaires pour la résolution numérique des problèmes de la physique mathématique (en russe), Novosibirsk (1966) et la bibliographie de ce livre.

DEUXIÈME THÈSE

Propositions données par la Faculté

Vu et approuvé

Paris, le 26 Juin 1969

Le Doyen de la Faculté des Sciences
de Paris

Marc. ZAMANSKY

Vu
et permis d'imprimer

Le Recteur de l'Académie de Paris

Jean ROCHE

